Motion artifact detection and correction in functional near-infrared spectroscopy: a new hybrid method based on spline interpolation method and Savitzky–Golay filtering

Sahar Jahani
Seyed K. Setarehdan
David A. Boas
Meryem A. Yücel

SPIE.

Motion artifact detection and correction in functional near-infrared spectroscopy: a new hybrid method based on spline interpolation method and Savitzky–Golay filtering

Sahar Jahani,[a,b] **Seyed K. Setarehdan,**[b] **David A. Boas,**[a,c] **and Meryem A. Yücel**[a,*]
[a]MGH/HST Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts, United States
[b]University of Tehran, College of Engineering, School of Electrical and Computer Engineering, Control and Intelligent Processing Center of Excellence, Tehran, Iran
[c]Boston University, Neurophotonics Center, Biomedical Engineering, Boston, Massachusetts, United States

Abstract. Motion artifact contamination in near-infrared spectroscopy (NIRS) data has become an important challenge in realizing the full potential of NIRS for real-life applications. Various motion correction algorithms have been used to alleviate the effect of motion artifacts on the estimation of the hemodynamic response function. While smoothing methods, such as wavelet filtering, are excellent in removing motion-induced sharp spikes, the baseline shifts in the signal remain after this type of filtering. Methods, such as spline interpolation, on the other hand, can properly correct baseline shifts; however, they leave residual high-frequency spikes. We propose a hybrid method that takes advantage of different correction algorithms. This method first identifies the baseline shifts and corrects them using a spline interpolation method or targeted principal component analysis. The remaining spikes, on the other hand, are corrected by smoothing methods: Savitzky–Golay (SG) filtering or robust locally weighted regression and smoothing. We have compared our new approach with the existing correction algorithms in terms of hemodynamic response function estimation using the following metrics: mean-squared error, peak-to-peak error (E_p), Pearson's correlation (R^2), and the area under the receiver operator characteristic curve. We found that spline-SG hybrid method provides reasonable improvements in all these metrics with a relatively short computational time. The dataset and the code used in this study are made available online for the use of all interested researchers. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.NPh.5.1.015003]

1 Introduction

Over the past 20 years, functional near-infrared spectroscopy (fNIRS) has gained much attention as an inexpensive and portable imaging technique that can noninvasively monitor brain function.[1–4] The technique is relatively robust to motion artifacts when compared to other imaging modalities, such as fMRI, and thus is a suitable technique for the research and real-life applications that involve excessive motion, such as walking, music performance, dancing, verbal communication or the study of infants and children.[5–7] Nonetheless, motion-induced artifacts still arise in the fNIRS signal during such tasks, and thus, the correction of motion artifacts is still an active research area.

fNIRS measurements consist of a source that emits light through the tissue and a detector that receives light that backscatters from the tissue.[1,4] Any head or skin movement can cause decoupling between the optode and the scalp, which results in abrupt changes in the amplitude of the received signal. These effects may cause different types of motion artifacts, such as high-frequency spikes or baseline shifts in the intensity

signal—the latter happens in cases, where the optode settles on a different location after the motion. Besides, slow head movements can also cause low-frequency motion artifacts.[8,9] Therefore, motion artifacts have a broad range of frequencies, which cannot be easily corrected just by frequency filtering without affecting the hemodynamic response function (HRF) estimation.

Ideally, motion artifacts should be avoided in the first place using various methods such as by stabilizing the optodes by gluing them on the scalp[10] or using a mechanical mounting structure to carry the weight of the optodes.[11,12] Auxiliary measurements, such as accelerometer, can also be used to regress out the motion artifacts from the fNIRS signal during postprocessing.[13–15] However, these methods are not always available in the experimental setup. Thus, various motion detection and correction techniques have been proposed.[16–20]

Motion artifacts are usually characterized by rapid signal changes larger in magnitude than the hemodynamic changes. Motion detection methods generally rely on user-defined input parameters to set a threshold for motion artifacts. By applying user-defined thresholds to various aspects of the signal,

*Address all correspondence to: Meryem A. Yücel, E-mail: mayucel@nmr.mgh.harvard.edu

such as moving standard deviation time series,[19] changes in absolute signal amplitude, or changes in the standard deviation of the data,[20] one can obtain the temporal epochs that are most likely contaminated by motion artifacts. Relying on user inputs introduces subjectivity into the detection of motion artifacts. Moreover, ideal parameters may change from one dataset to another. Ideally, the detection algorithm should objectively determine a threshold obtained directly from the signal itself based on deviations due to real physiological variations.

Motion correction algorithms used in the fNIRS community are rather broad. The simplest approach is to discard the artifact polluted trials from the analysis. This approach may not be ideal in cases where there are only a small number of trials.[21] Other widely used motion correction algorithms, which do not require any additional measurements, are principal component analysis (PCA),[22] targeted PCA (tPCA),[20] wavelet filtering,[18] correlation-based signal improvement (CBSI),[23] and movement artifact reduction algorithm (MARA).[19] PCA is very efficient when motion is the main source of variance in the signal but it requires multiple channels and also tends to overcorrect the signal.[22] To avoid overcorrection, tPCA was developed to only correct the preidentified epochs with motion artifacts.[20] The CBSI method assumes that HbO and HbR are always negatively correlated and become positively correlated only when there is a motion artifact. However, this assumption may not always hold, for example, during certain developmental stages or under abnormal brain physiology. The MARA method, which is also known as spline interpolation method, models the motions using spline interpolation and then subtracts them from the original signal. It is a fast and simple approach that corrects only motion artifact segments and is better at correcting baseline shifts than correcting high-frequency spikes. The method relies on the correct identification of motion artifact segments.[19] Wavelet filtering is based on decomposing the data into wavelet coefficients.[18] The outliers in the distribution of wavelet coefficients are assumed to reveal the presence of an artifact. Wavelet filtering is very effective in removing spikes but not in removing baseline shifts or low-frequency oscillations. In addition to wavelet filtering, there are some other denoising methods, such as Savitzky–Golay (SG) filtering and robust locally weighted regression and smoothing (Rloess). The SG method is a least square smoothing filter,[24] which substitutes each value of the signal series with a new value obtained from polynomial fitting to the successive subsets of adjacent data points. It has been previously used to denoise fNIRS signal.[25,26] Rloess combines weighted linear least squares regression with the flexibility of nonlinear regression.[27] In several fNIRS studies, Rloess has been used as a smoothing filter in data preprocessing.[28–30]

An ideal motion artifact method should have both an objective motion detection algorithm based on the deviations from real physiological variations, such as heart beat, and also should be able to deal with different types of motion artifacts, i.e., baseline shifts, high-frequency spikes, and slow motions. Each method mentioned already has its own strengths and weaknesses. In this work, we aimed to come up with an optimum approach that gathers the strengths of the existing methodologies for motion detection and correction. The new hybrid method first detects the motion artifacts based on the deviation from heartbeat variations and then corrects them based on their type, e.g., spline interpolation method for correcting baseline shifts followed by SG denoising algorithm for spike correction. To compare our proposed method with other methods, we added

a synthetic HRF to the raw signal and then recovered the HRF after the application of each motion correction algorithm. The extracted HRFs were then compared to the true HRF by the following metrics: mean-squared error (MSE), peak-to-peak error (E_p), Pearson's correlation (R^2), and the area under the receiver operator characteristic (ROC) curve (AUC).

2 Materials and Methods

The two datasets and the code used in this paper are made available online at Ref. 31 and the motion correction algorithms are implemented in the developer's version of the HOMER2 software, which is downloadable at Ref. 32. Please see the details of the datasets and the algorithms as follows.

2.1 Combined Spline Interpolation and Savitzky–Golay Filtering Method

A schematic diagram of the processing steps for the proposed method is shown in Fig. 1. The details of each block are described in the following sections.

2.1.1 Motion artifact detection algorithm

The intensity signal (I) is first low-pass filtered with a cut-off frequency of 2 Hz (I_{lpf}). Then, a Sobel filter ([$-1\,0\,1$]),
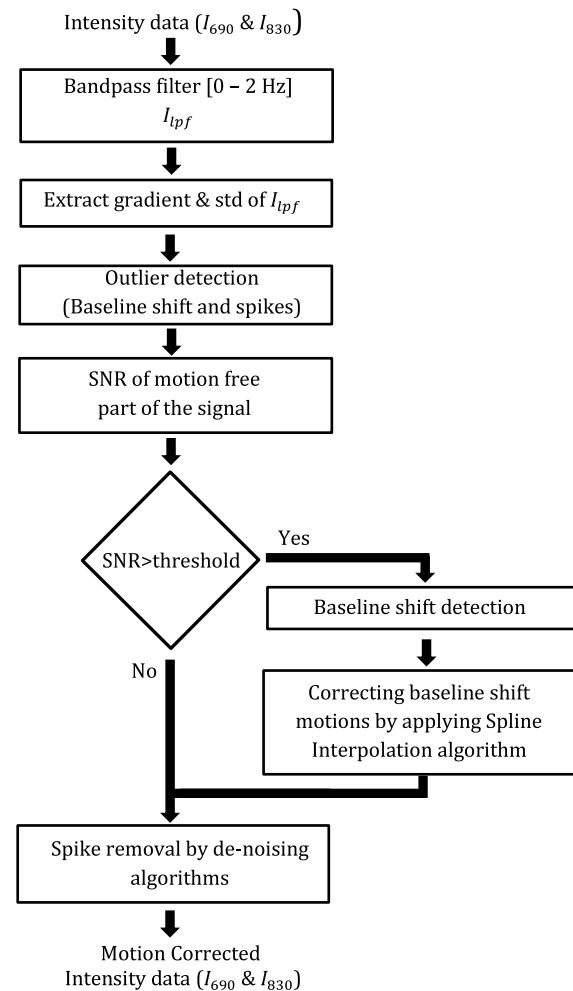

Fig. 1 The block diagram of the proposed method, I_{lpf} represents the low pass filtered intensity signal (I).

which is a discrete differentiation operator,[33] is used to compute the approximation gradient of I_{lpf}:

$$G = \text{convolution}([-1 \ 0 \ 1], I_{lpf}).$$

As motions create outliers in the gradient signal (G), we can identify them by finding outliers in G. The outliers are found by dividing the data into three equal parts. The values that divide each part are called the first, second, and third quantiles (Q1, Q2, and Q3). Q1 and Q3 are, respectively, the middle value in the first and second half of the rank-ordered G dataset. Q2 is the median value in the set. The interquartile range (IQR) is equal to Q3 minus Q1. The outliers are observations that fall below Q1 − 1.5*IQR or above Q3 + 1.5*IQR. Also, the standard deviation of the signal is extracted by sliding a window with a length of 1 s. Then, the outliers in the standard deviation of the signal are also detected by the same procedure applied on gradient. The unions of all detected outliers are considered as motion artifact.

2.1.2 *Detection of baseline shifts*

After finding all of the motion artifacts in the data including the baseline shifts, high-frequency spikes, and slow motions, as described already, we extracted the baseline shifts among them as follows. We first obtained the amplitude variation in the motion-free part of the signal (obtained using the algorithm in Sec. 2.1.1) by sliding a window with a length of 0.5 s, approximately half of the cardiac cycle period, in order to detect the maximum amplitude variation due to heart rate. Then, the maximum amplitude change in the motion-free signal is set as the threshold for finding baseline shifts. In other words, the detected motions higher than the heart beat oscillations are considered as baselines shifts. The baseline shifts detected are corrected by using spline interpolation method, as described in Sec. 2.1.4. After correcting the baseline shifts, the data with the remaining high-frequency spikes are fed to the SG (or Rloess) denoising algorithm. Also, spike type motions longer than 0.5 s (slow motion) are corrected by spline interpolation method as the denoising methods cannot remove them perfectly.

2.1.3 *Signal-to-noise ratio*

When the signal-to-noise ratio (SNR) of the data is poor, motion detection algorithms fail to find the motion artifacts and thus applying spline interpolation method to correct the baseline shifts results in new baseline shifts. To avoid this problem, we first calculated the SNR of the motion-free part of the signal (S) and then applied a predefined threshold (SNR = 3) on the SNR. We have chosen this threshold as we were able to extract the heart beat variation in the signal when the SNR of the signal was higher than 3. While the signals with high SNR (SNR > 3) are fed to the spline interpolation algorithm followed by SG denoising, the signals with low SNR (SNR < 3) are corrected by SG denoising only. The SNR was calculated as follows:

$$\text{SNR} = \text{mean}(S)/\text{std}(S).$$

2.1.4 *Spline interpolation method*

In this method, the motion artifact epoch is modeled by a cubic spline interpolation and the result is then subtracted from the original signal to get the denoised segment. Since the spline interpolation creates different signal levels between the original and the denoised signal, each segment needs to be shifted by a value defined with respect to its mean value and the mean value of the previous segment to reconstruct the whole time series. The accuracy of the interpolation depends on choosing a proper value for the parameter p in the range of [0, 1], which defines the degree of spline function. For $p = 1$, the interpolation will be a cubic spline, while for $p = 0$, the interpolation will be least-squares straight line. In this study, we used $p = 0.99$, which has been suggested in previous studies.[19,21] We used the hmrMotionArtifactbyChannel function in HOMER2 to detect motion contaminated parts of the signal[34] (parameters used for the first dataset: SDthresh = 20, AMPthresh = 5, tMotion = 0.5 s, tMask = 0.9 s and for the second dataset: SDthresh = 30, AMPthresh = 5, tMotion = 0.5 s, tMask = 0.9 s).

2.1.5 *Savitzky–Golay filtering*

SG filtering, also known as a digital polynomial filter or least square smoothing filter, is a digital smoothing filter that substitutes each value of the signal series with a new value, which is obtained from a polynomial fitting to the successive subset of adjacent data points.[24] The fitting is performed by the linear least squares fitting to $2n + 1$ neighboring points, while n can be equal or greater than the order of the polynomial. The more neighbors used in the averaging process, the smoother the signal becomes. Least squares smoothing suppresses noise while keeping signal information. We used $n = 300$ ($2n + 1 = 601$) for SG (note that our data were sampled at a rate of 50 samples/s). The parameter $2n + 1$ should be less than the length of HRF; otherwise, it can smooth out the HRF itself.

2.2 *Wavelet Filtering*

In wavelet filtering, the discrete wavelet transform is used to decompose the signal into multiple levels of scaling and wavelet coefficients.[18] The level of decomposition is based on the logarithm of the signal length. As the spikes have high-frequency content in comparison to the hemodynamic response, the wavelet coefficients higher than level 4 are considered as high-frequency components. Therefore, by finding outliers in levels higher than 4 and setting those to 0 before the inverse discrete wavelet transform should remove the spikes. In this study, the parameter IQR = 1.5 is used for finding the outliers in each wavelet decomposition level.

2.3 *Robust Locally-Weighted Regression and Smoothing Scatterplots*

Rloess is a nonparametric regression technique that combines much of the simplicity of linear least squares regression with the flexibility of nonlinear regression (second-degree polynomial model) models in a k-nearest-neighbor-based metamodel.[27] In this method, the fitted value x_k, $k = 1, \ldots, n$ for each x_i, $i = 1, \ldots, n$ (where n is the length of data) is the value of a polynomial fit to the data by using weighted least squares. The weight of the least squares depends on the inverse of the distance between x_i and x_k. Hence, the weight for x_i is large if x_i is close to x_k and is small if it is not. A zero weight is assigned to the data outside six mean absolute deviations. This robust fitting guards against outlier points, which are distorting the smoothing procedure.[27] The best value of the span for

Rloess is a value that removes the spike while keeping the frequency of the HRF (<0.5 Hz). In this study, we used $R = 0.02$ for the Rloess method.

2.4 Targeted Principal Component Analysis

PCA uses a linear orthogonal transformation technique to convert the signal time-course among all N channels into a set of N uncorrelated variables. The eigenvalues and eigenvectors are extracted from the covariance matrix of measurements. As the motions constitute the majority of the variation in the signal, the larger eigenvalues are assumed to represent the variance caused by motion artifacts. So, by removing the first k eigenvalues, it will result in removing a specific percentage of the variance in the signal.[21,22]

Applying PCA to a selected set of data points that are assumed to have motion artifact will reduce the risk of eliminating the physiological variations in the motion-free part of the signal.[20] The prerequisite for this method is the identification of motion artifacts. The signal change within a certain time period (tMotion) is considered a motion artifact if the amplitude or standard deviation in that time period exceeds a predefined threshold (AMPthresh and SDthresh, respectively). After removing 97% of the variance in the motion-detected part of the signal, the signal is stitched back to the original signal by a shifting procedure. This process was repeated two more times to detect any remaining motion artifacts. The parameters of this algorithm were set to their effective value as reported in previous studies.[20] Motion detection parameters were the same as spline interpolation method: for the first dataset: SDthresh = 20, AMPthresh = 5, tMotion = 0.5 s, tMask = 0.9 s and for the second dataset: SDthresh = 30, AMPthresh = 5, tMotion = 0.5 s, tMask = 0.9 s. The motion detection in the tPCA-SG method was done with the approach introduced in Sec. 2.1.1.

2.5 Correlation-Based Signal Improvement

The CBSI method is based on the hypothesis that HbO and HbR are negatively correlated, and they only become positively correlated when a motion artifact occurs.[23] So, this method is based on two assumptions: (1) the true HbO and HbR (x_0 and y_0) should be negatively correlated ($x_0 = -\beta y_0$) and (2) the correlation between the true HbO and the motion artifact (F) should be close to 0. The measured HbO and HbR signal (x and y) are modeled as follows:

$$x = x_0 + \frac{\text{std}(x)}{\text{std}(y)} * F + \text{Noise} \quad y = y_0 + F + \text{Noise},$$

where std is the standard deviation of the signal. Then, by applying the two aforementioned assumptions on the model and considering $\alpha = \beta$, the following corrected signal can be obtained:

$$x_0 = \frac{1}{2}\left(x - \frac{\text{std}(x)}{\text{std}(y)} * y\right), \quad y_0 = -\frac{\text{std}(y)}{\text{std}(x)} * x_0.$$

Please note that the corrected signals "x_0" and "y_0" do not represent HbO and HbR and are simply a linear combination of measured HbO and HbR signals.
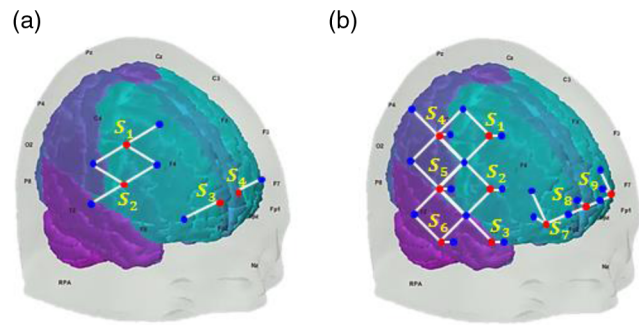

Fig. 2 The schematic view of the optical probe for (a) dataset I and (b) dataset II. The red and blue circles indicate the position of sources and detectors, respectively. The sources are named by yellow letters. The white lines indicate source–detector pairs.

2.6 fNIRS Datasets

In this study, two separate datasets are used to evaluate the performance of the proposed method. The study was approved by Massachusetts General Hospital (MGH). The methods were carried out in accordance with the guidelines and regulations of the Institutional Review Board of MGH. A written consent form is obtained from each subject.

The first dataset was obtained from seven healthy adults by using a TechEn CW6 system (Medford, Massachusetts). The subjects were asked to perform the following movements to create motion artifacts throughout the 6-min recording: reading aloud, nodding their head up and down, nodding sideways, twisting upper body right, twisting upper body left, shaking head rapidly from side to side and raising their eyebrows. The probe equipped with 6 laser sources and 10 long-distance detectors resulted in 14 channels of data (30 mm separation; Fig. 2). For further details about the dataset, please refer Ref. 10. Dataset II was acquired from five healthy subjects during resting state by using a TechEn CW6 system (Medford). The data contained a minimum number of artifacts to test the impact of motion correction on data with few motion artifacts. The probe consisted of 15 sources, 18 long separation detectors, and 14 short separation detectors that are located 30 and 8 mm away from the sources, respectively. The sampling frequency for both dataset is 50 Hz.

2.7 Processing Stream for Hemodynamic Response Function Estimation

A synthetic HRF was generated in the raw NIRS data by introducing a signal change of 1% from baseline for the 690 nm signal and 2% for 830 nm signal, resulting in a 0.6 μM increase in HbO and 0.2 μM decrease in HbR with a pathlength factor of 6 for both wavelengths.[35,36] The simulated HRF was a gamma function, which peaks around 6 s and lasts for 16 s. The HRF interstimulus interval ranged from 5 to 10 s, which provided 17 to 19 stimuli during a 6.5-min recording. About 25 different random stimulus onsets were created and convolved with the true HRF. The noisy channels lower than 80 dB and higher than 140 dB were pruned from further processing. (Please note that for our NIRS system, the SNR is typically less than 10 when the signal is below 80 dB, and the signal saturates above 140 dB.) The intensity data were then converted into optical density (OD) and then the proposed motion artifact detection algorithm was applied to the OD time-series. The identified baseline shifts and slow motion artifacts were
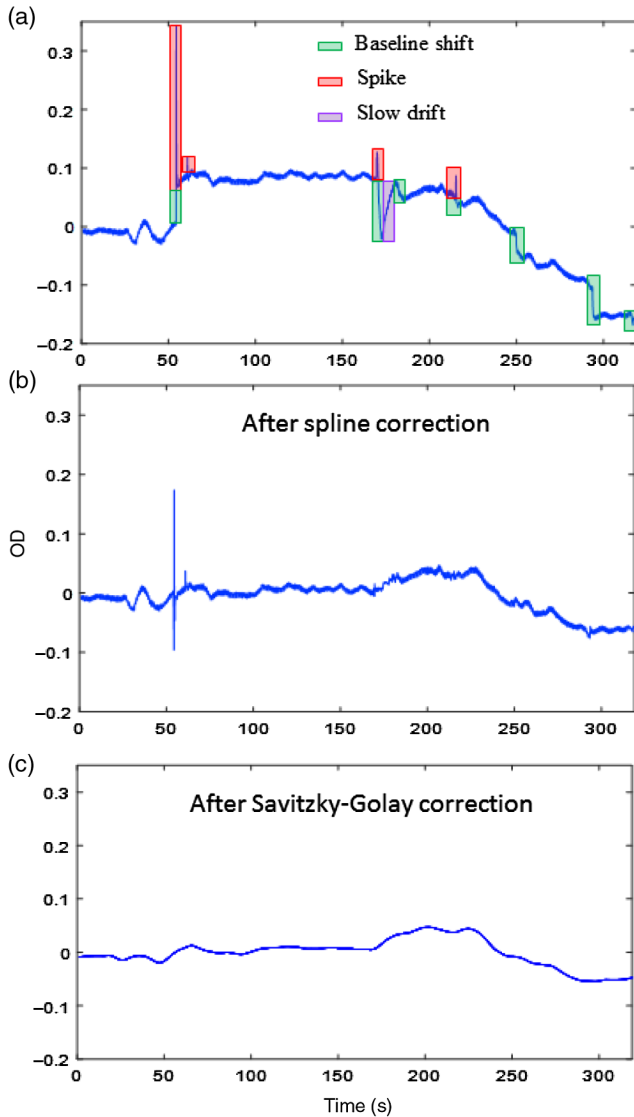
(a)

(b)
After spline correction

(c)
After Savitzky-Golay correction

Fig. 3 (a) An example of an OD change time course at 690 nm contaminated by three types of motions: baseline shifts (green), high-frequency spikes (pink), and slow drifts (violet). (b) The baseline shifts and slow drift motions are corrected by spline interpolation method and (c) spikes are corrected by SG denoising.

corrected by using the hmrMotionArtifactSpline function in HOMER2[34] and the spikes were removed using SG (MATLAB function: sgolayfilt) for the spline-SG method or were removed using Rloess (MATLAB function: smooth) filters for the spline-Rloess method. The data were then low-pass filtered with a cut-off frequency of 0.5 Hz. Finally, the HRF was estimated using a general linear model (GLM) with short separation regression. The GLM method uses the least square method to estimate the weights of consecutive temporal basis functions. We used Gaussian basis functions with a standard deviation of 0.5 s and separated means by 0.5 s as done in our previous studies.[37] To correct the drift, we modeled the baseline drift with a third order polynomial fit.

2.8 Evaluation Criteria

The following four metrics were used to compare the true HRF with the recovered HRF after the application of each motion correction algorithm: (1) R^2, (2) MSE, (3) E_p, and (4) AUC. The R^2 metric is obtained using the corr function in MATLAB (MathWorks Inc., Natick, Massachusetts). The MSE and E_p are defined as follows:

$$\text{MSE} = \frac{\|\widehat{\text{HRF}} - \text{HRF}\|^2}{\|\text{HRF}\|^2} \times 100,$$

$$E_p = \frac{\|\widehat{\text{HRF}}_p - \text{HRF}_p\|^2}{\|\text{HRF}_p\|^2} \times 100,$$

where HRF and $\widehat{\text{HRF}}$ represent the true and estimated HRF, respectively, with the subscript p designating the peak amplitude. The MSE and E_p were normalized by dividing with the corresponding amplitude of the true HRF. ROC curves were obtained by plotting the true-positive rate against the false-positive rate using the t-statistic of the HRF estimation as the threshold for detection. This is done by taking the mean of the t-values obtained from each of the 25 different stimulus paradigms.

Paired student's t-tests were used to evaluate statistically significant differences between motion correction algorithms. For statistical analysis, multiple comparison correction is applied using Benjamini–Hochberg method with a false discovery rate of 0.05.[38]
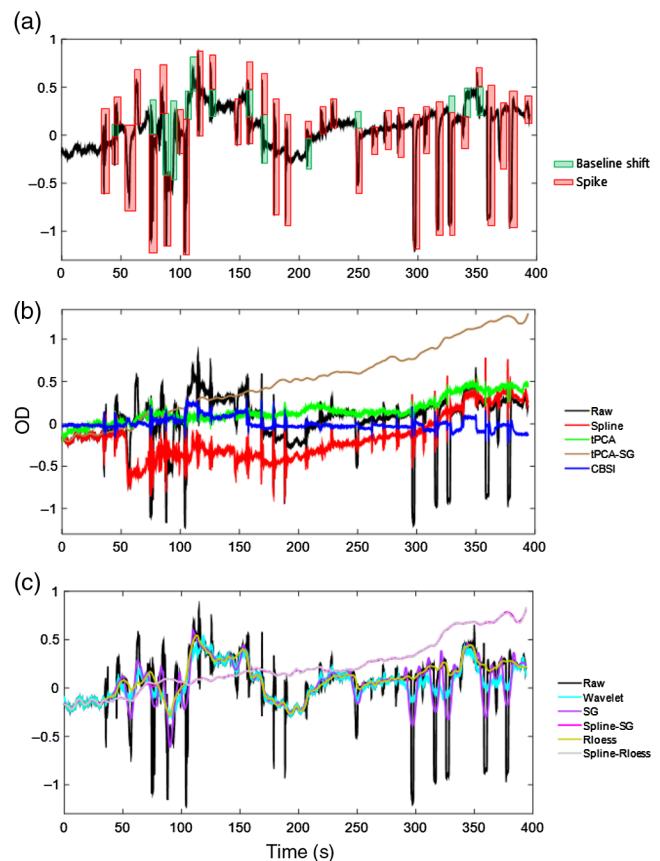
(a)

(b)

(c)

Fig. 4 (a) OD change at 690 nm from a single channel with no synthetic HRF with baseline shifts (green), high-frequency spikes (pink), zoomed out data (b and c) before correction (black) and after correction by spline (red), tPCA (green), tPCA-SG (brown), wavelet (cyan), CBSI (dark blue), SG (violet), spline-SG (magenta), Rloess (dark gold), and spline-Rloess (silver).

3 Results

Figure 3 depicts an example of motion detection and correction by the proposed spline-SG algorithm. As can be seen in Fig. 3(a), the selected channel is contaminated by all three types of motion artifacts, i.e., baseline shifts, spikes, and slow drift artifacts highlighted with different color coding. The corrected signal after applying spline interpolation method on baseline shifts and the denoised signal after applying the SG filtering algorithm are shown in Figs. 3(b) and 3(c), respectively.

To show the effect of different motion correction techniques, a highly motion-contaminated signal is shown before and after being corrected with spline interpolation, tPCA, tPCA-SG, CBSI, wavelet, SG, spline-SG, Rloess, and spline-Rloess

motion correction algorithms (Fig. 4). As shown in Figs. 4(b) and 4(c), wavelet smoothing and the SG methods corrected the high-frequency spikes; however, they were ineffective with baseline shifts. On the other hand, the tPCA, CBSI, and spline methods left residuals at the beginning and the end of motion artifact epochs. The amplitude of this residual was highest after applying spline interpolation method. The Rloess method showed a better performance than wavelet and SG by completely removing the spikes without propagating baseline shifts. The hybrid methods, spline-SG, tPCA-SG, and spline-Rloess show the best performance among them all in terms of correcting baseline shifts and high frequency spikes. Using spline or tPCA first has removed all baseline shifts and
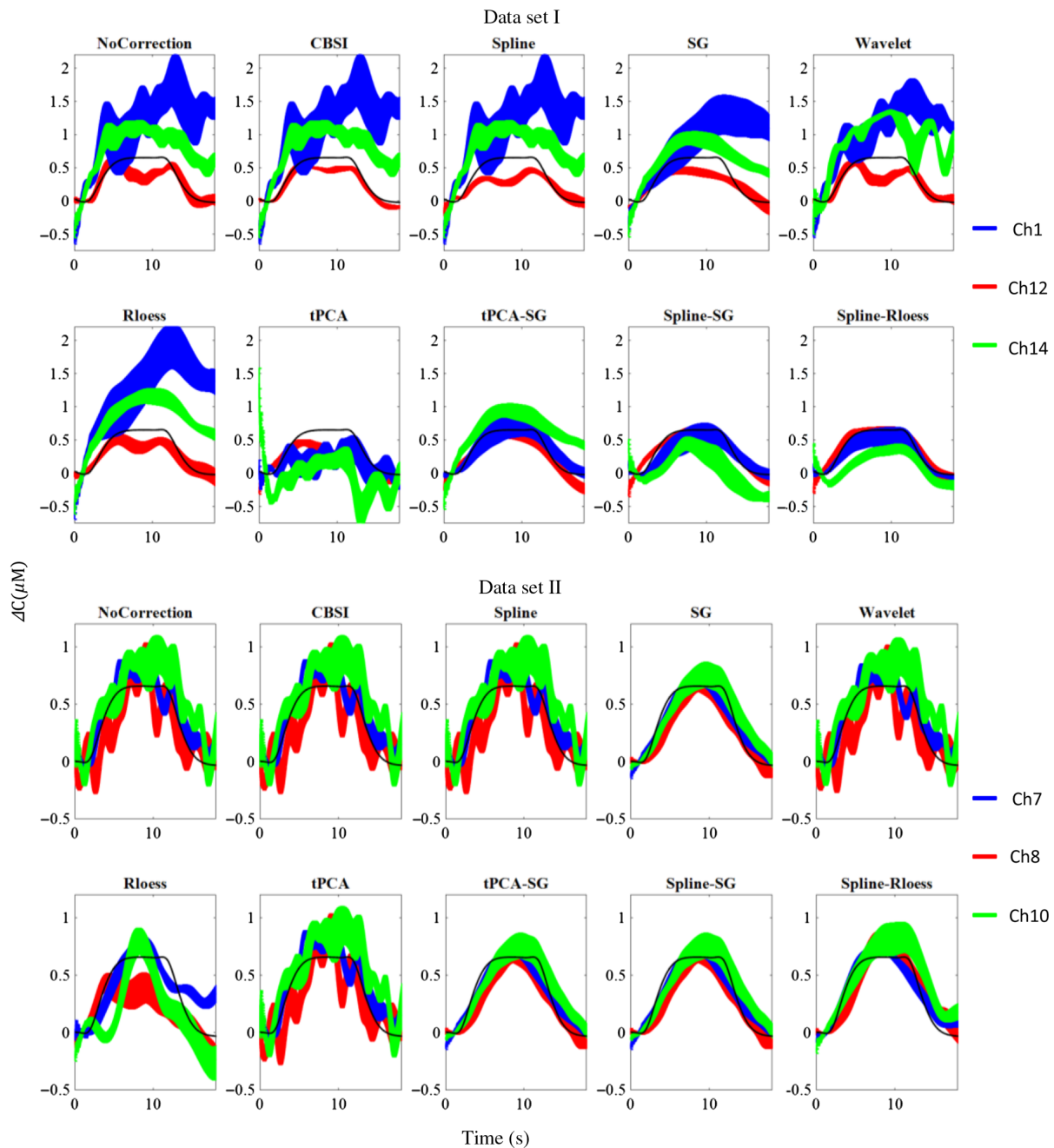

Fig. 5 The true and recovered HRFs are shown for different channels from datasets I and II. The shaded areas represent standard deviation.

slow drifts and the remaining spikes were successfully removed by SG or Rloess.

3.1 Performance of the Motion Correction Algorithms for Improving HRF Estimation

To compare the proposed spline-SG method with the existing methods, synthetic HRFs were first added to the raw intensity signal and then recovered after each motion correction algorithm. The HRF recovered at different channels is shown as an example in Fig. 5. The first two rows show channels from dataset I, which is highly contaminated by motion artifacts. For the specific examples, spline, wavelet, tPCA, Rloess, and CBSI methods resulted in HRFs with larger deviations from the true HRF while the amplitude of the recovered HRF after spline-SG, spline-Rloess, and tPCA-SG was closer to the true HRF. For dataset II, which has fewer motion artifacts compared to dataset I, most of the methods showed reasonable results (Fig. 5). The hybrid methods were also more effective in this dataset by producing smoother HRFs with acceptable amplitude and correlation compared to the other methods.

We have evaluated the performance of each method using five different metrics: (1) the coefficient of determination (square of Pearson's correlation coefficient, R^2), (2) the MSE, (3) the E_p, (4) the AUC, and (5) the total processing time. The MSE, E_p, and R^2 values are averaged over all subjects, channels, and different stimulus paradigms. The AUC values are measured over the subjects. The results of MSE, E_p, R^2, and AUC for datasets I and II are summarized in Tables 1 and 2 and Figs. 6 and 7. The statistics are provided in Table 3.

For dataset I, all motion correction algorithms improved the recovered HRF in comparison to no correction (Table 1 and Fig. 6). While spline-SG, spline-Rloess, and tPCA-SG methods performed better than the rest in terms of E_p and MSE, Rloess, spline-SG, and spline-Rloess have shown the best correlation with the true HRF. Rloess, spline-Rloess, spline-SG, and CBSI methods produced higher AUC values for the ROC curves

compared to the rest. The wavelet, Rloess, and spline-Rloess methods had the longest processing time.

Overall, the mean E_p and MSE were lower and R^2 was higher in dataset II compared to dataset I, as this second dataset involved fewer motion artifacts than the first one. In this dataset, the denoising algorithms, such as wavelet, SG, tPCA-SG, spline-SG, and spline-Rloess, resulted in better E_p, MSE, and R^2 than spline and tPCA (Table 2 and Fig. 7). SG and spline-SG produced the lowest E_p and MSE, while the CBSI method had the highest correlation and AUC compared to the other algorithms. For both datasets, spline-SG was the optimum method among them all with considerable improvement in all metrics and a relatively short processing time.

For dataset I, the improvement in E_p, MSE, and R^2 obtained after spline-Rloess was statistically better compared to all other algorithms [paired t-test, p-value < 0.01; see Table 3(a), second row]. The spline-SG was second to it in terms of performance [Table 3(a), first row]. For the second dataset, SG alone and spline-SG methods outperformed the rest of the methods [paired t-test, p-value < 0.01; Table 3(b)]. There was no significant difference between mentioned algorithms in terms of AUC.

4 Discussion

In this study, we sought the optimum algorithm for motion detection and correction for fNIRS data. In order to do this, we have introduced a motion detection algorithm, which is solely based on the variation in the NIRS signal and thus does not require input from the user. The detection algorithm defines the baseline shift motion artifacts by comparing the deviation in the signal with real physiological fluctuations, namely the heartbeat. We have combined this motion detection algorithm with an ideal combination of existing methods, such as tPCA-SG, spline-SG, and spline-Rloess. In order to assess the performance of the proposed new approach, we have compared it with the wavelet, spline, tPCA, and CBSI methods, which are the most common methods in fNIRS data processing.

Table 1 MSE, E_p, R^2, and AUC values for dataset I from different motion reduction algorithms and no correction. The processing time is calculated for 14 channels (6.6 min of recording with 50-Hz sampling rate) on a 3.4-Hz CPU running windows 7. The bold values are the best values among all algorithms. The proposed method is highlighted in italics.

Dataset I	MSE $\times 10^4$	$E_p \times 10^4$	R^2	AUC	Processing time (s)
No correction	2.26 ± 0.38	12.58 ± 2.34	0.68 ± 0.02	0.87 ± 0.03	—
Spline	1.59 ± 0.45	8.75 ± 2.71	0.73 ± 0.03	0.89 ± 0.02	0.8
Wavelet	1.22 ± 0.28	6.96 ± 1.71	0.73 ± 0.03	0.88 ± 0.03	135
CBSI	2.23 ± 0.40	13.48 ± 2.60	0.77 ± 0.02	$\mathbf{0.91 \pm 0.03}$	0.2
tPCA	1.14 ± 0.27	7.30 ± 1.98	0.71 ± 0.03	0.85 ± 0.03	1.8
tPCA-SG	0.89 ± 0.21	5.50 ± 1.35	0.74 ± 0.02	0.85 ± 0.03	14
SG	1.50 ± 0.31	9.32 ± 2.28	0.74 ± 0.03	0.87 ± 0.02	$\mathbf{0.08}$
Spline-SG	*0.65 ± 0.17*	*4.20 ± 1.18*	*0.79 ± 0.02*	*0.89 ± 0.02*	*6*
Rloess	1.30 ± 0.33	8.19 ± 2.27	0.79 ± 0.02	0.90 ± 0.02	764
Spline-Rloess	$\mathbf{0.60 \pm 0.16}$	$\mathbf{3.90 \pm 1.13}$	$\mathbf{0.80 \pm 0.02}$	0.89 ± 0.02	767

Table 2 MSE, E_p, R^2, and AUC values for dataset II from different motion reduction algorithms and no correction. The processing time is calculated for 36 channels (6.6 min of recording with 50-Hz sampling rate) on a 3.4 Hz CPU running windows 7. The bold values are the best values among all algorithms. The proposed method is highlighted in italics.

Dataset II	MSE $\times 10^4$	$E_p \times 10^4$	R^2	AUC	Processing time (s)
No correction	0.75 ± 0.11	4.34 ± 0.73	0.79 ± 0.02	0.88 ± 0.05	—
Spline	0.73 ± 0.11	4.24 ± 0.72	0.79 ± 0.02	**0.89 ± 0.05**	2
Wavelet	0.55 ± 0.08	2.97 ± 0.49	0.81 ± 0.01	0.88 ± 0.05	529
CBSI	0.54 ± 0.09	3.39 ± 0.56	**0.84 ± 0.01**	0.84 ± 0.01	0.6
tPCA	0.76 ± 0.12	4.31 ± 0.72	0.74 ± 0.02	0.85 ± 0.06	10
tPCA-SG	0.51 ± 0.07	3.10 ± 0.54	0.82 ± 0.01	0.86 ± 0.07	46
SG	0.45 ± 0.06	2.53 ± 0.41	0.83 ± 0.01	0.88 ± 0.05	**0.1**
Spline-SG	***0.44 ± 0.06***	***2.52 ± 0.41***	*0.83 ± 0.01*	***0.89 ± 0.05***	*16*
Rloess	0.72 ± 0.10	4.26 ± 0.72	0.79 ± 0.02	0.87 ± 0.05	1800
Spline-Rloess	0.56 ± 0.08	3.36 ± 0.59	0.82 ± 0.01	**0.89 ± 0.05**	1800

Fig. 6 (a) MSE, (b) E_p, (c) R^2, and (d) AUC values for estimated HRF before (no correction) and after correcting the motions using tPCA (light blue), spline (dark green), wavelet (orange), CBSI (dark blue), SG (dark cyan), Rloess (violet), tPCA-SG (khaki), spline-SG (pink), and spline-Rloess (light green) for dataset I. The bars represent the mean and the error bars symbolize the standard error over all subjects and stimulus paradigms.
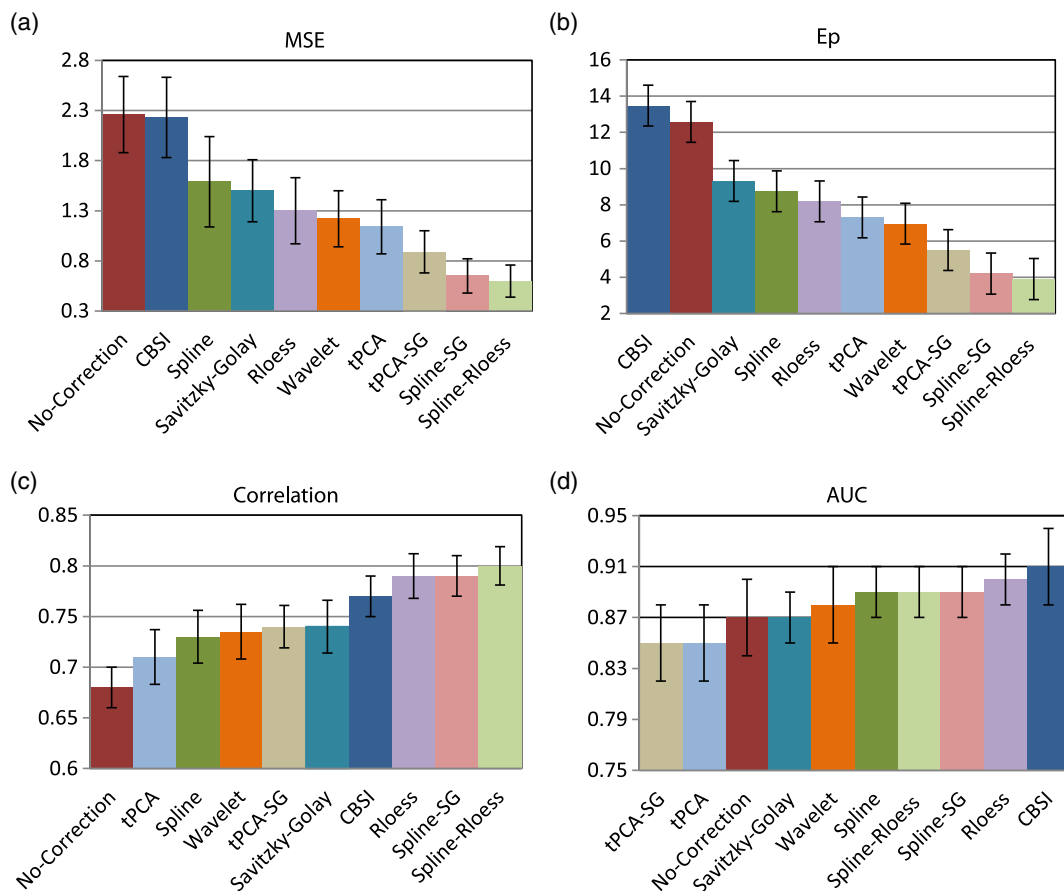
Fig. 7 (a) MSE, (b) E_p, (c) R^2, and (d) AUC values for estimated HRF before (no correction) and after correcting the motions using tPCA (light blue), spline (dark green), wavelet (orange), CBSI (dark blue), SG (dark cyan), Rloess (violet), tPCA-SG (khaki), spline-SG (pink), and spline-Rloess (light green) for dataset II. The bars represent the mean and the error bars symbolize the standard error over all subjects and stimulus paradigms.
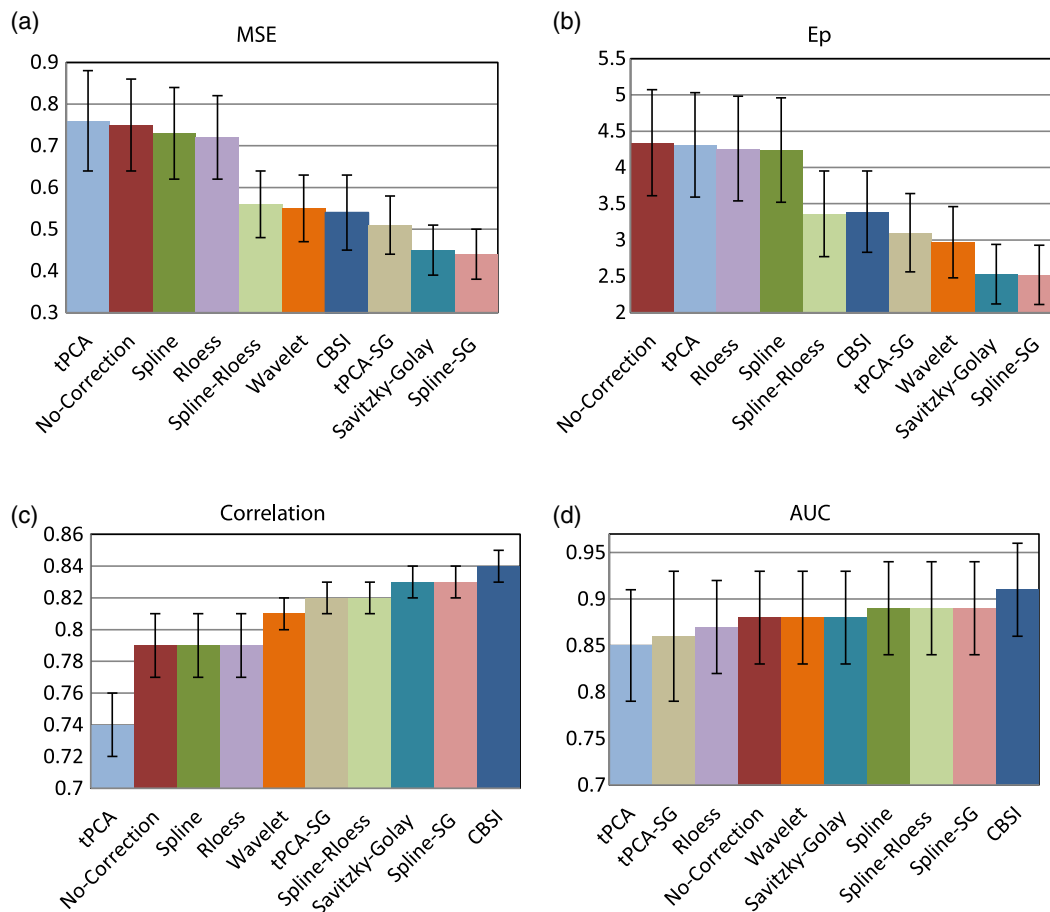
The wavelet, Rloess, and SG methods are denoising methods, which are very powerful in removing high-frequency motion artifacts, but not in correcting baseline shifts or slow motion artifacts. The Rloess method performed much better compared to SG and wavelet in terms of following the trend of the original signal in time periods, where there are baseline shifts. The method along with wavelet, however, has the drawback of a high computational cost. The proposed approach, spline-SG, deals well with both baseline shifts as well as high-frequency spikes by combining the strengths of each method. The method first applies a spline interpolation method to correct the baseline shifts. The signal is then passed to the SG denoising algorithm for the correction of high-frequency spikes. This way, SG does not have to deal with baseline shifts and thus produces better results. Moreover, the hybrid spline-SG method does not have the disadvantage of long processing times as the Rloess method or spline-wavelet combination (not shown here), which makes it the optimum method among the existing correction algorithms. The smoothing parameter for the spline-SG method though should be carefully chosen. It should allow the suppression of high-frequency spikes while keeping the slower variations in the signal that can correspond to a hemodynamic response.

Our performance metrics, i.e., the R^2, MSE, E_p, and AUC for the ROC curve, were based on the estimation of the HRF, which is synthetically added to the raw signal. We also added the processing time as a metric to evaluate the practicality of the method. Our results show that the spline-SG approach produced statistically better results in HRF recovery in terms of E_p, MSE, R^2, and AUC compared to the other methods for both datasets. Considering the good performance in HRF estimation metrics and the short processing time, we suggest the use of the spline-SG method for motion correction in fNIRS data analysis.

5 Conclusion

Motion artifact correction is often an essential preprocessing step during HRF estimation, particularly during experimental paradigms in which subjects are engaged in natural behaviors. In this work, we introduced a new way of motion artifact detection and also combined the strengths of the existing methods in order to optimally deal with all types of motion artifacts. The proposed approach, spline-SG, provides a suitable solution to the motion artifact removal problem in fNIRS studies by combining the spline's powerful baseline shift correction with the powerful spike correction of SG. The results of this study suggest that the proposed algorithm performs better than the

Table 3 The p-values (nonbold) signify significant improvement in E_p, MSE, or R^2 using the method in the row over the method in the column. The p-values in bold, on the other hand, show superiority of the method in the column over the method in the row. All p-values are corrected for multiple comparisons.

(a) Dataset I

E_p / MSE / Corr	Spline-SG	Spline-Rloess	tPCA-SG	SG	Rloess	Wavelet	tPCA	CBSI	Spline	No-correction
Spline-SG	—	**<0.01**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	**<0.01**	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	**<0.01**	<0.01	<0.01	—	<0.01	<0.01	—	<0.01	<0.01
Spline-Rloess	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	<0.01	<0.01	—	<0.01	<0.01	<0.01	<0.01	<0.01
tPCA-SG	—	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	—	—	**<0.01**	—	<0.01	**<0.01**	—	<0.01
SG	—	—	—	—	**<0.01**	**<0.01**	**<0.01**	<0.01	<0.01	<0.01
	—	—	—	—	**<0.01**	—	**<0.01**	<0.01	<0.01	<0.01
	—	—	—	—	**<0.01**	—	<0.01	**<0.01**	<0.01	<0.01
Rloess	—	—	—	—	—	**<0.01**	—	<0.01	—	<0.01
	—	—	—	—	—	—	—	<0.01	<0.01	<0.01
	—	—	—	—	—	<0.01	<0.01	<0.01	<0.01	<0.01

(b) Dataset II

E_p / MSE / Corr	Spline-SG	Spline-Rloess	tPCA-SG	SG	Rloess	Wavelet	tPCA	CBSI	Spline	No-correction
Spline-SG	—	<0.01	<0.01	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	<0.01	<0.01	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	**<0.01**	<0.01	<0.01
Spline-Rloess	—	—	<0.01	<0.01	<0.01	**<0.01**	<0.01	—	<0.01	<0.01
	—	—	<0.01	<0.01	<0.01	—	<0.01	—	<0.01	<0.01
	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	**<0.01**	<0.01	<0.01
tPCA-SG	—	—	—	<0.01	<0.01	—	<0.01	—	<0.01	<0.01
	—	—	—	<0.01	<0.01	—	<0.01	—	<0.01	<0.01
	—	—	—	<0.01	<0.01	<0.01	<0.01	**<0.01**	<0.01	<0.01
SG	—	—	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	—	—	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
	—	—	—	—	<0.01	<0.01	<0.01	**<0.01**	<0.01	<0.01
Rloess	—	—	—	—	—	**<0.01**	—	**<0.01**	—	—
	—	—	—	—	—	**<0.01**	—	**<0.01**	—	—
	—	—	—	—	—	**<0.01**	<0.01	**<0.01**	—	—

existing methods with a considerably short processing time. Thus, we recommend the use of spline-SG method for the correction of motion artifacts in the NIRS signal.

Disclosures

Acknowledgments

References

1. D. A. Boas et al., "Twenty years of functional near-infrared spectroscopy: introduction for the special issue," *Neuroimage* **85**, 1–5 (2014).
2. A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends Neurosci.* **20**(10), 435–442 (1997).
3. D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy," *Neuroimage* **23**(Suppl. 1), S275–S288 (2004).
4. S. Lloyd-Fox, A. Blasi, and C. E. Elwell, "Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy," *Neurosci. Biobehav. Rev.* **34**(3), 269–284 (2010).
5. T. Farroni et al., "Infant cortex responds to other humans from shortly after birth," *Sci. Rep.* **3**, 2851 (2013).
6. H. Karim et al., "Functional near-infrared spectroscopy (fNIRS) of brain function during active balancing using a video game system," *Gait Posture* **35**(3), 367–372 (2012).
7. L. A. Tuscan et al., "Exploring frontal asymmetry using functional near-infrared spectroscopy: a preliminary study of the effects of social anxiety during interaction and performance tasks," *Brain Imaging Behav.* **7**(2), 140–153 (2013).
8. S. Brigadoi 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).
9. F. Scholkmann, A. J. Metz, and M. Wolf, "Measuring tissue hemodynamics and oxygenation by continuous-wave functional near-infrared spectroscopy—how robust are the different calculation methods against movement artifacts?" *Physiol. Meas.* **35**(4), 717–734 (2014).
10. M. A. Yücel et al., "Reducing motion artifacts for long-term clinical NIRS monitoring using collodion-fixed prism-based optical fibers," *Neuroimage* **85**, 192–201 (2014).
11. S. M. Coyle et al., "Brain-computer interface using a simplified functional near-infrared spectroscopy system," *J. Neural Eng.* **4**(3), 219–226 (2007).
12. P. Giacometti and S. G. Diamond, "Compliant head probe for positioning electroencephalography electrodes and near-infrared spectroscopy optodes," *J. Biomed. Opt.* **18**(2), 027005 (2013).
13. S. G. Diamond et al., "Dynamic physiological modeling for functional diffuse optical tomography," *Neuroimage* **30**(1), 88–101 (2006).
14. A. J. Metz et al., "A new approach for automatic removal of movement artifacts in near-infrared spectroscopy time series by means of acceleration data," *Algorithms* **8**(4), 1052–1075 (2015).
15. J. Virtanen et al., "Accelerometer-based method for correcting signal baseline changes caused by motion artifacts in medical near-infrared spectroscopy," *J. Biomed. Opt.* **16**(8), 087005 (2011).
16. M. Izzetoglu et al., "Motion artifact cancellation in NIR spectroscopy using Wiener filtering," *IEEE Trans. Biomed. Eng.* **52**(5), 934–938 (2005).
17. A. M. Chiarelli et al., "A kurtosis-based wavelet algorithm for motion artifact correction of fNIRS data," *Neuroimage* **112**, 128–137 (2015).
18. B. Molavi and G. A. Dumont, "Wavelet-based motion artifact removal for functional near-infrared spectroscopy," *Physiol. Meas.* **33**(2), 259–270 (2012).
19. F. Scholkmann et al., "How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation," *Physiol. Meas.* **31**(5), 649–662 (2010).
20. M. A. Yucel et al., "Targeted principle component analysis: a new motion artifact correction approach for near-infrared spectroscopy," *J. Innov. Opt. Health Sci.* **7**(2), 1350066 (2014).
21. R. J. Cooper et al., "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy," *Front. Neurosci.* **6**, 1–10 (2012).
22. Y. Zhang et al., "Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging," *J. Biomed. Opt.* **10**(1), 011014 (2005).
23. X. Cui, S. Bray, and A. L. Reiss, "Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics," *Neuroimage* **49**(4), 3039–3046 (2010).
24. A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.* **36**(8), 1627–1639 (1964).
25. N. T. Hai et al., "Temporal hemodynamic classification of two hands tapping using functional near-infrared spectroscopy," *Front. Hum. Neurosci.* **7**, 516 (2013).
26. A. Vrana et al., "Cortical sensorimotor processing of painful pressure in patients with chronic lower back pain—an optical neuroimaging study using fNIRS," *Front. Hum. Neurosci.* **10**, 578 (2016).
27. W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Am. Stat. Assoc.* **74**(368), 829–836 (1979).
28. Z. Zhang et al., "The comparisons of cerebral hemodynamics induced by obstructive sleep apnea with arousal and periodic limb movement with arousal: a pilot NIRS study," *Front. Neurosci.* **10**, 403 (2016).
29. L. Corvaglia et al., "Bolus vs. continuous feeding: effects on splanchnic and cerebral tissue oxygenation in healthy preterm infants," *Pediatr. Res.* **76**(1), 81–85 (2014).
30. T. D. Yanowitz et al., "Variability in cerebral oxygen delivery is reduced in premature neonates exposed to chorioamnionitis," *Pediatr. Res.* **59**(2), 299–304 (2006).
31. https://www.nitrc.org/projects/fnirsdata/.
32. HOMER2, MGH-Martinos Center for Biomedical Imaging, https://www.nitrc.org/projects/homer2/.
33. K. Engel, "Sobel operator," in *Real-time Volume Graphics*, K. Engel et al., Eds., pp. 112–114, A K Peters, Ltd., Wellesley, Massachusetts (2006).
34. T. J. Huppert et al., "HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain," *Appl. Opt.* **48**(10), D280–D298 (2009).
35. M. Cope and D. T. Delpy, "System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination," *Med. Biol. Eng. Comput.* **26**(3), 289–294 (1988).
36. D. T. Delpy et al., "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.* **33**(12), 1433–1442 (1988).
37. L. Gagnon et al., "Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling," *Neuroimage* **56**(3), 1362–1371 (2011).
38. Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. R. Stat. Soc. B* **57**(1), 289–300 (1995).

Biographies for the authors are not available.