ABSTRACT

Functional Near Infrared Spectroscopy (fNIRS) is an emerging neurological sensing technique applicable to optimizing human performance in transportation operations, such as commercial aviation. Cognitive state can be determined via pattern classification of functional activations measured with fNIRS. Operational application calls for further development of algorithms and filters for dynamic artifact removal. The concept of using the frequency domain phase shift signal to tune a Kalman filter is introduced to improve the quality of fNIRS signals in real-time. Hemoglobin concentration and phase shift traces were simulated for four different types of motion artifact to demonstrate the filter. Unwanted signal was reduced by at least 43%, and the contrast of the filtered oxygenated hemoglobin signal was increased by more than 100% overall. This filtering method is a good candidate for qualifying fNIRS signals in real time without auxiliary sensors.

Keywords: cognition, fNIRS, in-task monitoring, Kalman filter, real time, signal processing

1 INTRODUCTION

Functional Near Infrared Spectroscopy (fNIRS) is an emerging neurological sensing technique applicable to optimizing human performance in transportation operations, such as commercial aviation. FNIRS quantifies hemoglobin concentration ([Hb]) changes in the brain based on optical intensity measurements.
Hemodynamic activations can be detected across the whole head. These ambulatory, non-invasive measurements allow in-task monitoring. Cognitive state is determined via pattern classification of the functional activations. Thus, we can watch the activations within the brain of a pilot during the safety-critical task of flying. Importantly, this is distinct from the use of vigilance tests which the pilot could undergo only while breaking from the task at hand. Such "fit-to-fly" tests may be passed due to short-term increases in attentional effort (Sarter, 2006) which may not be sustainable for long durations. During monitoring, information regarding cognitive state could be used to trigger appropriate risk mitigations in real time via changes in flight automation and information display.

Our work continues to focus on improving techniques for applying fNIRS to in-task operator characterization. Optical sensing may improve upon the existing use of other modalities for operator characterization, especially through combination with them. Examples include electroencephalography and other physiological measures used for Augmented Cognition (Schmorrow, 2007), operator performance research (Schnell, 2004), and crew cognition research (Pope, 1995).

FNIRS quantifies [Hb] changes with time based on optical intensity measurements of light that has scattered through the outer layers of the cortex beneath the optical probe. FNIRS measures the same hemodynamic changes as the functional Magnetic Resonance Imaging (fMRI) Blood Oxygen Level Dependent signal, with lower spatial but improved temporal resolution. fNIRS provides a direct connection to the wealth of information in the fMRI literature. We believe this improves the outlook for sensing more complicated states.

2 MOTIVATION

FNIRS works well in the laboratory, and it has been used in many neuroscientific research studies, including real-time classification of state using an extended Kalman filter and known stimulus timing (Abdelnour, 2009). However, implementation for monitoring outside the laboratory requires techniques that do not rely on known stimulus timing or block averaging. Also, no standard for the removal of physiological noise and motion artifact yet exists. Many opportunities remain for real-time applications (Zhang 2009) and for improving robustness and reliability. Motion artifact can be significant, and is likely to occur in operational environments. Field application calls for further development of algorithms and filters for the automation of bad channel detection and dynamic artifact removal. Such advances in operational use will benefit use in clinical outpatient scenarios.

Frequency domain (FD) instruments for fNIRS measure radio-frequency-modulated signal intensity amplitude and offset, plus phase shift of the detected optical intensity signal relative to that of the source. The phase shift data provide a direct indication of the coupling noise associated with signal detection and thus the quantification of [Hb] at that time for that channel. Thus we explore the use of this phase measurement to drive an adaptive filter for the removal of motion artifact in real-time.
Principally we are interested in avoiding human error to improve commercial aviation safety by identifying the loss of attentional engagement using fNIRS signals. The attentional monitoring system under development (Harrivel, 2010) could be adapted for the detection of other cognitive states by changing the location of the optodes and adjusting the classifier training data and parameters. Our previous work has demonstrated improvements in the comfort of optical head probes to enable monitoring for over an hour (Harrivel, 2009), and has shown on average 70% accuracy for real time classification of attentional state using Support Vector Machines with training data free of artifact (Harrivel, 2011). To improve hemodynamic activation measurements and classification accuracy, we continue to implement existing cutting edge techniques while exploring novel headgear and data processing methods to provide such clean data in real time during operational use in the field. Here we introduce a filter based on the FD phase measurement for real-time signal quality detection and improvement. This new filtering technique is well-suited to improving the quality of fNIRS signals in any application where motion is an inherent part of the task to be performed during monitoring.

3 RATIONALE

3.1 Optical rationale

Optical intensity is attenuated by absorption and scattering along the optical path length (the light’s path through the head). The phase is sensitive to motion because of its dependence on the source-detector separation and the index of refraction of the tissue through which the light travels. Both affect the optical path length. An increase in the actual optical path length causes greater absorption. This can lead to overestimation of the change in [Hb] if the path length is assumed to be constant. High variability of the FD phase shift can indicate inconsistent coupling at the optode-scalp interface, which changes the separation distance and possibly exposes the detector to ambient light (which has no consistent phase shift with respect to the source). Shifts in the phase also can indicate motion of a probe along the scalp surface. As the probe senses different volumes of tissue, changes in intensity are caused by changes in the optical scattering and absorptive properties of the tissue. This also contributes to probe relocation errors and inter-subject variability.

If the phase and [Hb] changes are correlated, it is likely the [Hb] changes are due to motion-induced changes in path length or absorption. This is because the increased phase indicates a larger path length, which causes more absorption and lowers the detected intensity, which finally increases the calculated [Hb]. If the phase is steady, it is likely the calculated [Hb] changes are due to activations of interest, systemic physiology or detector gain changes. Thus, the phase can be monitored to automate signal removal or alert the user to reset the probe.

This filter assumes the phase does not change significantly with physiological activation. Neuron swelling or the influx of new hematocytes may affect the scattering and thus the optical path length and phase, but not on time scales
accessible to FD instrumentation.

This filtering method is most appropriate for the Frequency Domain Multiple Distance (FDMD) method where FD instrumentation is used to calculate optical absorption and scattering properties so absolute [Hb] values can be determined (Fantini, 1999). In this case, the filter would be applied to the measured optical intensity, and anti-correlation between the phase and the intensity could be used. This can be explored in future work. Alternately, in cases where relative measures suffice while probe footprints on the head must be small and the signals must be reliable, it can be used to filter the relative [Hb] changes calculated using the Modified Beer Lambert Law. The phase signal used should be for the wavelength that matches the hemoglobin species of the trace being filtered. This is the case presented here.

3.2 The Kalman filter and its suitability

Kalman filtering (Kalman, 1960) is a technique for estimating the state of a linear discrete dynamical system. The Kalman filtering implementation used for this work assumes that the unknown true value $x_k$ and measurement value $z_k$ of a system at time $k$ is determined by the model

$$x_k = Ax_{k-1} + w_{k-1},$$
$$z_k = Hx_k + v_k,$$

where $A$ and $H$ are known linear operators. $A$ and $H$ were each assumed to be an identity operator for this work. In this way, we are assuming that the process under study is simply the contamination of the data by additive noise. The variables $w_k$ and $v_k$ are process and measurement noise, respectively. The distributions of these two noise sources are assumed to be

$$w_k \sim N(0, Q) \text{ and } v_k \sim N(0, R),$$

where $N(\mu, \sigma^2)$ is a normal probability distribution with mean $\mu$ and variance $\sigma^2$. The Kalman filter produces estimates, by means of the explicit recurrence relations, of the data

$$\hat{x}_k = \hat{x}_{k-1} + \frac{\hat{x}_{k-1} + Q}{\hat{x}_{k-1} + Q + F}\left(x_k - \hat{x}_{k-1}\right),$$

and the variance of the filter

$$\hat{p}_k = \left(1 - \frac{\hat{x}_{k-1} + Q}{\hat{x}_{k-1} + Q + F}\right)(\hat{p}_{k-1} + Q),$$

beginning with the initial condition $x_0 = z_0$ and $p_0 = 1$. Because of the explicit nature of these computations, no iterative solver or optimization procedures are required for filter implementation. Furthermore, the computed estimates $\hat{x}_k$ are optimal in the mean squared error sense. More specifically,

$$\hat{x}_k = \arg\min_{\hat{x}_k} E((x_k - \hat{x}_k)^2),$$

for each time $k$; where $E$ is the mathematical expectation operator. The optimal properties of the Kalman filter estimates along with the simplicity of its implementation (Zarchan, 2005) make it an excellent choice for the smoothing of fNIRS data traces.
Values for the variances of the process and measurement noise ($Q$ and $R$) must be determined before the Kalman filter can be implemented; since they are required in the calculation of each recurrence relation (each $k$). These values can be interpreted as tuning parameters for the Kalman filter. A larger $R$ weights the data estimate and distrusts the measurement to produce a smoother output, while a smaller $R$ follows the measured signal more closely. $Q$ can be increased to allow more dynamic changes in the estimate of the system.

We propose that the phase channel $\phi$ associated with a [Hb] channel $x^*$ is a natural choice of information for computing an appropriate value for $R$ at each time.

## 4 METHODS

Synthetic data were created to mimic [Hb] and phase shift traces with a data rate of 6.25Hz for 400s to demonstrate the feasibility of this filter. Four [Hb] time series were simulated: relative changes in both oxygenated and deoxygenated (reduced) hemoglobin species concentration ([HbO] and [HbR]) for two source-detector separations, which we label close and far. A close source can be used to measure systemic physiological contributions to the signal of interest (biological “noise”) on a per-sensor basis (Zhang, 2009) by interrogating shallow tissue (not cortex). The far source interrogates both superficial and cortical tissue.

The input trace to be filtered was created by adding non-dynamic physiological signals for cardiac at 1Hz, respiration at 0.25 Hz, and Meyer waves at 0.1Hz to the close source traces. Functional activations, and smaller physiological contributions, were added to the far source traces. Random noise was added to all traces. In figure 3, the input [Hb] trace is shown (top, black) after within-species subtraction of the close source signal from the far source signal, followed by subtraction of [HbR] from [HbO]. Functional activations were simulated with gamma-variant functions as in Abdelnour, et al. (2009) eq. 3, with the [HbR] amplitude being $-1/4$ of [HbO].

Four different types of motion artifact were simulated. These are outlined in Table 1. No motion artifact is simulated from 0s to 50s. Artifact from probe slip (across skin with no air gap) was simulated between 50s and 100s by allowing the phase to vary randomly over tens of degrees. Artifact from decoupling (an air gap at the detector interface) was simulated between 100s and 200s with phase varying over hundreds of degrees. Artifact from a probe bump was simulated with a spike in the phase at 225s, while artifact from a bump and relocation was simulated with a spike and a step increase of 10 degrees at 325s. Phase changes on these orders of magnitude do occur in real data. [Hb] signal which covaried with the phase was generated for the slip and the relocation bump artifacts.

In the case where a noisy signal could appear to be a [Hb] change (50s – 200s), filter performance is quantified as percentage of signal mean reduction. In the cases where motion noise obscures the activation of interest, performance is quantified as contrast-to-noise ratio (CNR) improvement between the unfiltered and the filtered signal. CNR is defined here as mean activation signal minus mean reference signal (0s to 50s), divided by the square root of the sum of the variances in those signals.
4.1 Mathematical Implementation of the Kalman filter

To implement the Kalman filter, the measurement noise variance, $R$, can be set to a function, possibly non-linear, of the variance in the phase over a window of time prior to that instance. The use of the phase differentiates our method from other implementations of Kalman filtering for motion artifact removal (Izzetoglu, 2010). Here, the tuning parameter $R$ was set to the variance of the phase in a rolling window of 1s prior to the current instance, plus the correlation of the phase with the [Hb] trace being filtered during 20s prior (which turns on at 20s). The contributions of the variance and correlation were linearly scaled. The absolute value of the correlation was added if it was greater than a threshold of 0.3. Thus only high correlations impact $R$. This threshold has not been optimized, and can be raised for less aggressive filtering.

4.2 Parameter Selection

Parameters were selected empirically for this pilot study to maximize signal retention and artifact rejection. Different scale factors and rolling time window sizes were explored for both the variance and the correlation. The scale factors were selected to bring the magnitude of $R$ into the useful range of 0 to 5. $R$ should be small (~0) to allow desired changes in signal to be retained, and large to remove artifact. $R$ on the order of 2 or more was found to remove artifact effectively (as in figure 3, 100s – 150 s).

Increasing window size increases signal plateau due to filter turn off delay, as seen in figure 3 at 340s. It is due to correlation contributing to $R$ even after it has passed in time. This may be tolerable for conservative investigators who wish to err on the side of signal removal. Shorter windows may be less able to detect artifact. Figures 1 and 2 show improved rejection (signal mean reduction) of the probe slip (50s - 100s) and the probe relocation (325s) artifacts attained by allowing the correlation to contribute to $R$.

$Q$ was set to 1e-4 or 1e-5 as noted in the figures. If larger, high frequency spikes are not removed. If smaller, desired signal changes are attenuated. The initial state estimate was set to $z_0 = 0$, and the initial variance of the filter was set to $p_0 = 1$. All of these parameters are available for adjustment and optimization, with the objective of maximizing the rejection of artifact and the retention of good signal.

In real application, optimal parameters may depend on source-detector separation. For example, the signal to noise ratio of the phase itself depends on the source-detector separation (Dehaes, 2011). Ideally, the [HbO] close and far, and [HbR] close and far traces should be filtered with their own parameters.
5 RESULTS

After filtering, the mean of the unwanted peak between 50s and 100s is reduced by 43%. That between 100s and 200s is completely removed. The reference signal mean is reduced by 9%. CNR is increased by more than 100% overall. The results are listed in Table 1. The filter output trace is shown in figure 3 (negative, top).
Figure 3 Calculated R (bottom) used to filter the [Hb] trace (top, filter output in negative). Both the phase variance and the correlation are used to calculate R. Q was set to 1e-5.

Table 1 Phase-based adaptive Kalman filter performance by motion type

<table>
<thead>
<tr>
<th>Time</th>
<th>Motion simulated</th>
<th>Phase shift signal feature</th>
<th>Filter Performance Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>0s-50s</td>
<td>none (reference)</td>
<td>none (reference)</td>
<td>• CNR increase: 125%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• signal mean reduction: 9%</td>
</tr>
<tr>
<td>50s-100s</td>
<td>probe slip</td>
<td>slow changes correlated with [Hb]</td>
<td>• unwanted signal mean reduction: 43%</td>
</tr>
<tr>
<td>100s-200s</td>
<td>probe decoupling</td>
<td>very high variance</td>
<td>• unwanted signal mean reduction: 100%</td>
</tr>
<tr>
<td>200s-300s</td>
<td>transient probe bump</td>
<td>spike at 225s</td>
<td>• CNR increase: 144%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• signal mean reduction: 12%</td>
</tr>
<tr>
<td>300s-400s</td>
<td>probe bump, relocation</td>
<td>spike and step increase at 325s, correlated with [Hb]</td>
<td>• CNR increase: 204%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• signal mean increase: 11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• late filter turn off extends activation by 10s</td>
</tr>
</tbody>
</table>

6 DISCUSSION

An advantage of this filtering method is that it only comes on when needed according to actual conditions without adding a channel or auxiliary sensor. The phase provides an objective indication of whether [Hb] changes are due to motion without relying only on the [Hb] trace. Also, the phase can simply be used to
indicate data quality for complete removal during post-processing.

A disadvantage is the introduction of time delays. The filter can only change the estimate so quickly, and delays can be on the order of seconds when the filter is on (high R). Filter turnoff delay is also a problem, as discussed in section 4.2 and as evidenced by the artificial increase in signal mean between 300s and 400s. Q may be increased and time windows shortened to reduce delays.

Other methods for real time artifact removal include simply smoothing, using other signal correlations, or using a moving standard deviation. With smoothing, unwanted signal is not rejected, and the temporal resolution advantage of fNIRS is eroded in all instances. The Kalman filter can be adjusted to quickly follow non-noisy data by instance. However, smoothing does not introduce delay.

Cui, et al. (2010) present a successful method for removing motion artifact from fNIRS signals which takes advantage of the anti-correlation between [HbO] and [HbR] inherent in functional activation. However, the phase-based filter removes motion artifact from both channels intended to measure functional activations and channels for systemic physiology, which lack neural activation.

A moving standard deviation measurement on the [Hb] signal itself has been used to detect spikes for removal (Scholkman, 2010). However, mechanical changes can degrade the signal via poor coupling resulting in a relatively flat [Hb] signal. This would not trigger a moving standard deviation detection algorithm.

7 CONCLUSIONS AND FUTURE WORK

The phase is a good candidate for qualifying fNIRS signals in real time, enabling automated bad-channel detection and motion artifact removal without auxiliary sensors. Successful future work would increase the amount of motion which can be tolerated while obtaining useful fNIRS signals. This real-time filtering technique relies upon frequency-domain instrumentation, but is well-suited to improving the quality of systemic physiology and functional activation measurements with fNIRS in any operational environment where motion is an inherent part of the task being performed.

Further parametric study is needed to minimize delays and maximize the performance of the filter for data collected with human subjects. The investigation of non-linear functions of the phase variance and the correlation between the phase and the [Hb] is planned. The filter should be tried on the measured optical intensity with the FDMD method. Human subject studies to quantify the motion of the probe with respect to the head during the phase shift measurement should be undertaken. Ultimately, cognitive state classification could be performed with this filter to determine the effect on classification accuracy.

ACKNOWLEDGMENTS

The authors would like to acknowledge NASA’s Aviation Safety Program and the Biomedical Engineering Department at the University of Michigan.
REFERENCES


