Published in final edited form as:

Ann Biomed Eng. 2019 February ; 47(2): 646–658. doi:10.1007/s10439-018-02135-7.

Quantifying Movement in Preterm Infants Using Photoplethysmography

Ian Zuzarte¹, Premananda Indic², Dagmar Sternad³, and David Paydarfar⁴

¹Department of Bioengineering, Northeastern University, Boston, MA

²Department of Electrical Engineering, University of Texas, Tyler, TX

³Departments of Biology, Electrical and Computer Engineering, and Physics, Northeastern University, Boston, MA

⁴Department of Neurology, Dell Medical School, and Institute for Computational Engineering and Sciences, The University of Texas, Austin, TX

Abstract

Long-term recordings of movement in preterm infants might reveal important clinical information. However, measurement of movement is limited because of time-consuming and subjective analysis of video or reluctance to attach additional sensors to the infant. We evaluated whether photoplethysmogram (PPG), routinely used for oximetry in preterm infants in the neonatal intensive care unit (NICU), can provide reliable long-term measurements of movement. In 18 infants (mean post-conceptual age: 31.10 weeks, range 29 to 34.29 weeks), we designed and tested a wavelet-based algorithm that detects movement signals from the PPG. The algorithm's performance was optimized relative to subjective assessments of movement using video and accelerometers attached to two limbs and force sensors embedded within the mattress (5 infants, 3 raters). We then applied the optimized algorithm to infants receiving routine care in the NICU without additional sensors. The algorithm revealed a decline in brief movements (<5s) with increasing post-conceptual age (13 infants, $r=-0.87$, $p<0.001$, post-conceptual age range 27.3 – 33.9 weeks). Our findings suggest that quantitative relationships between motor activity and clinical outcomes in preterm infants can be studied using routine photoplethysmography.

Keywords

Continuous wavelet transform; movement detection; motor development; preterm movement; photoplethysmography

1. Introduction

Infants born prematurely (<37 weeks gestational age) are at increased risk of motor, cognitive, social and behavioral abnormalities that can extend beyond childhood. Minor neurological dysfunction and more severe complications such as cerebral palsy and autistic

disorder are often associated with premature birth. It is therefore imperative to better understand the development of physiological mechanisms to preempt or moderate complications at later ages.

Movements in preterm infants have shown to be an important predictor for motor and cognitive abilities and their development. For example, the quality of motor activity of preterm infants early in life is linked to cognitive outcomes at school age. Periods of movement can also correlate with other physiological activities and adverse events in preterm infants, for example episodes of apnea. Therefore, early assessment of motor patterns in premature infants may be important for neonatal care delivery as well as therapeutic interventions that affect long-term outcomes.

To date, there is little knowledge about the normal developmental stages of motor activity in preterm infants. One reason for this lack of knowledge is that there are many constraints on the study of preterm infants because they are extremely fragile. Even though preterm infants are monitored continuously in the Neonatal Intensive Care Units (NICU), most assessments of motor activity have been limited to qualitative observations by clinicians at the bedside or from video recordings. Such evaluation is episodic, subjective, and requires expertise of the clinician performing the evaluation. The procedure is also time-consuming, susceptible to observer fatigue and attention, and suffers from low intra- and inter-observer repeatability. Automated movement recognition has been explored using video motion capture systems, which are costly and difficult to set up in a clinical environment. Other methods to capture movement include sensors attached to the infant, which have limited use because of the lack of validation studies, and the reluctance of approval of the sensors and additional wires for medical use.

Rather than adding more sensors to the infant, we have turned to a sensor that is routinely used in clinical practice, the pulse oximeter probe. This probe, placed on the infants' hand or foot, is an optical sensor that non-invasively measures arterial oxygen saturation and blood volume changes. Importantly, infant movements cause artifacts in the PPG by generating non-pulsatile mechanical perturbations of venous blood and perhaps other changes in optical coupling between the sensor and the skin. Previous studies on movement-related PPG signals have focused on mitigating its artifactual nature and improving estimation of pulse rate and SpO_2 . However, these techniques have not explored the usefulness of this movement "artifact" *per se* and how it can serve as a vital signal.

The goals of the present study were two-fold. First, we sought to develop a computational method that extracts the artifactual distortions of the photoplethysmogram caused by movement. We found that a wavelet-based algorithm can efficiently quantify the movement-related perturbations of the pulse signal. The algorithm's performance was optimized relative to subjective assessments of movement in 5 infants using video and accelerometers attached to two limbs and force sensors embedded within the mattress. Our second goal was to evaluate the clinical applicability of the optimized algorithm to infants receiving routine care in the NICU, using photo plethysmography without the aid of additional sensors. We found in an additional group of 13 infants that the algorithm revealed salient developmental changes in movement activity over a range of post-conceptual ages (PCA).

2. Materials and Methods

2.1. Human Subjects

Eighteen preterm infants were studied at the University of Massachusetts Memorial Healthcare NICU. All infants were spontaneously breathing room air. Infants with hydrocephalus, congenital defects, bronchopulmonary disease, and intraventricular hemorrhage higher than grade 2 were excluded. Eligible infants were identified to the investigators by the attending physician. Written informed consent was obtained from the infant's mother or legal guardian. The study was approved by the University of Massachusetts's Medical School Institutional Review Board for Human Subjects. Table 1 lists subject characteristics and study duration.

2.2. Data Acquisition

All subjects were studied in their NICU incubator with a bedside monitor (IntelliVue MP70, Philips Medical Systems, Andover, MA), which displayed electrocardiogram (ECG), pneumogram and photoplethysmogram (PPG) signals, blood oxygen saturation ($SpO_2\%$), heart and respiration rates. The PPG signal was acquired from the infant's hand or foot using a pulse oximeter probe (Masimo SET LNCS Neo, Masimo, Irvine, CA). A soft foam tapeless wrap (Posey wrap 6554, Posey Co., Arcadia, CA) was placed over the probe to secure it in place and prevent interference from ambient light and optical cross-talk between the sensors. Infants #1–5 were studied to develop and optimize our algorithm using additional sensors to detect movement. For these infants, the mattress was replaced by a specially-constructed mattress (Wyss Institute, Boston, MA) that had four strips of force-sensing resistors (FSR 408, Interlink Electronics, Camarillo, CA) embedded in the mattress foam (Fig. 1). Further, two tri-axial accelerometers (ADXL, Analog Devices, Norwood, MA) were attached to the infant, one at the same limb of the pulse oximeter sensor and one at another limb at the nurse's discretion. A video camera, placed in one corner of the incubator, continuously monitored the infant (Resolution: 320×240, Edmund Optics, Barrington, NJ). Data were recorded using the Vuelogger™ Patient Monitoring System (Wyss Institute, Boston, MA), which retrieved physiological information from the bedside monitors and time-synced them with signals from the video, the accelerometers, and force sensors in the mattress. The signals were sampled at the following rates: 500Hz (ECG, force-sensing resistors and accelerometers); 62.5Hz (pneumogram); 125Hz (PPG); 10fps (video). The data were streamed to the Vuelogger's hard-disk at the bedside and then were exported for analysis. Infants #6–18 were studied to determine if the optimized algorithm reveals important clinical information regarding the maturation of motor patterns in infants of different ages using routine prolonged continuous recordings of photo plethysmography. These infants received routine NICU care using their standard mattress and monitoring, and there were no additional movement sensors attached to their limbs.

2.3. Detection of Movement from Pulse Plethysmograph

The pulse oximeter is a device that measures oxygen saturation of arterial blood (SpO_2) and pulse rate from a source signal, the photoplethysmogram (PPG). Without movement, the waveform is highly periodic with the pulse. Movement disrupts the pulsatile nature of the PPG signal, producing non-stationary fluctuations. While these disruptions by movement are

generally considered artifacts that compromise the measurement of pulse and arterial oxygen saturation, we hypothesized that they can be used to measure the onset and duration of movements. Therefore, we designed an algorithm that can identify movement-generated artifacts in the PPG signal.

The algorithm used a wavelet transformation that converts a time signal to the time-frequency domain. For the discrete PPG signal x_n , we applied the continuous wavelet transform $W_n(\jmath)$ defined as the convolution of x_n with a scaled and translated version of a wavelet function ψ_0 :

$$W_n(\jmath) = \sum_{k=0}^{N-1} x_k \psi^* \left(\frac{(k-n)\delta t}{\jmath} \right) \quad n = 0, 1, 2, \dots, N-1 \quad (1)$$

where N is the number of data points in the time series, n is the localized time index, $\jmath$ is the wavelet scale and δt is the sampling rate of the time-series. By varying $\jmath$ and translating along n , the wavelet transform provides information about the variation in the PPG time-series at different scales and locations.

The scales were represented as fractional powers of two:

$$\jmath_j = \jmath_o 2^{p\delta j} \quad p = 0, 1, 2, \dots, J \quad (2)$$

Where $\jmath_o$ is the smallest resolvable scale and J determines the total number of scales. With the PPG time-series sampled at $\delta t = 125\text{Hz}$, we chose $\jmath_o = 2\delta t$, $\delta j = 1/32$ and $J = 256$. The term $\frac{(k-n)\delta t}{\jmath}$ represents a non-dimensional ‘time’ parameter and can be denoted as η . The $(*)$ in Eqn (1) indicates the complex conjugate of $\psi(\eta)$, which is the normalized representation of the mother wavelet function $\psi_0(\eta)$ to have unit energy and is given by:

$$\psi(\eta) = \left(\frac{\delta t}{\jmath} \right)^{\frac{1}{2}} \psi_0(\eta) \quad (3)$$

The choice of the appropriate $\psi_0(\eta)$ depends on the nature of the type of information to be extracted from the signal. The Morlet wavelet has the best frequency resolution, while Paul wavelet achieves the best time localization because the transform is less affected by edge effects. In previous work with the PPG signal we corroborated these differences between the two transforms and since the main aim of our study was to capture the onset of movement with temporal precision, we used the following Paul wavelet:

$$\psi_o(\eta) = \frac{2^m i^m m!}{\sqrt{\pi(2m)!}} (1 - i\eta)^{-(m+1)} \quad (4)$$

The order m controls the number of oscillations in the mother wavelet. Since, smaller values of m achieve better time resolution, we chose $m=4$. The result of the wavelet transform is a scalogram, which provides insight into the frequency distribution at every time sample of the signal.

In our previous study, we applied a similar wavelet transform to the entire PPG signal. This method had several limitations: 1) At any given point in time, detection of the movement onset required data from future time points that prevented application of the algorithm in real-time. 2) The large sample size N led to a poor accuracy of movement onset and offset detection. 3) Bradycardias reduced the frequency of the pulse waveform, which increased the power in the scalogram. The latter was a serious shortcoming, because its effects on the results were indistinguishable from those of a movement. The revisions of the algorithm developed here removed these limitations.

2.3.1. Real-Time Implementation with Optimal Temporal Resolution—The continuous wavelet transform of the PPG signals was computed in a moving window of N samples and shifted each time by a fixed number of samples. A small window size of $N=200$ samples (at a sampling rate of $125\text{Hz} = 1.6\text{s}$) was selected to improve the time localization of the wavelet transform (Figs. 2 and 3A, 3B). The wavelet power spectrum, defined as $|W_n(s)|^2$, was obtained from the continuous wavelet transform of the 200-point time series.

This gives the local measure of the PPG energy distribution at each scale s and time ($n=0,1,2,\dots,199$) (Figs. 3C, 3D). To obtain an estimate of movement at the last time point in the 200-point window, we used the information from the wavelet power spectrum from 25-time points towards the end of the window: $n=150-174$ (time = 3.6 to 3.8s in Fig. 3C and 9.6 to 9.8s in Fig. 3D). The last 25 points, $n=175-199$ (time = 3.8 to 4s in Fig. 3C and 9.8 to 10s in Fig. 3D) of the wavelet power spectrum were excluded, since the wavelet transformation of a finite-length time-series resulted in degraded detection of power at the edges of the window. The time-averaged wavelet power spectrum over the desired range (Figs. 3E, 3F) for the i^{th} window was then obtained as:

$$\bar{W}_i^2(s) = \frac{1}{25} \sum_{n=150}^{174} |W_n(s)|^2 \quad (5)$$

Movement artifacts in the PPG signal consisted of low-frequency interference ($<0.67\text{Hz}$), which translated to a Fourier period (λ) greater than 1.5s in the wavelet power spectrum. The relation between the equivalent Fourier period and wavelet scale s for the Paul wavelet is given by:

$$\lambda = \frac{4\pi}{2m+1} s \quad (6)$$

with the order of the wavelet, $m=4$; $\lambda=1.396s$. Using this relation, wavelet scales were converted into Fourier periods. The estimate of movement in each 200-point window was

taken from the maximum power in time-averaged power spectrum during the Fourier period greater than 1.5s.

$$Mov(i) = \max_{\lambda > 1.5} (\overline{W}_i^2(\lambda)) \quad (7)$$

The window of 200 points was shifted every 5 points (40ms) to maintain a good resolution of the estimated movement signal. As the PPG signal was sampled at 125Hz (8ms/sample), the estimated movement signal had a sampling rate of 25Hz (40ms/sample).

The different steps of the algorithm are shown in Fig. 3, using the two segments S1 and S2 marked in Fig. 2 as examples. Figs. 3A and 3B show the PPG segments during non-movement (S1) and movement periods (S2). Figs. 3C and 3D show the scalograms of the two segments obtained with the Paul wavelet. Figs. 3E and 3F display the average power from the 25-point window (samples 150–174). The dashed red line indicates the Fourier period threshold of 1.5s. As can be seen, the non-movement segment has its highest power around a period of 0.43s, which corresponds to the infant's heart rate at around 140bpm. The movement segment shows considerably more power at higher periods (>1.5s) corresponding to low-frequency fluctuations in the PPG. The maximum value of the average power at periods >1.5s was the estimated movement for each of the two PPG segments. The units of the estimated movement are the same as the units of power of the PPG signal. Since the PPG signal obtained from the NICU monitor was unitless, the estimated movement was expressed in arbitrary units (au). However, the PPG signal was bounded by a fixed range of values between 1024 and 3072. Therefore, based on this range, the estimated movement was between 0 and 115au. The algorithm can be implemented in real time with a small delay of 200ms, since at any point in time it only uses time-average power from the previous 200ms window. It did not utilize any information from data points obtained at a future time step.

To determine the onset and offset of movement, thresholds merely based on the amplitude of the estimated movement were insufficient, because the baseline levels differed across subjects. These inter-individual differences could be attributed to differences in peripheral perfusion and/or skin pigmentation. Therefore, thresholds had to be defined by a combination of amplitude, standard deviation and the slope of a preceding window (Table 2). We empirically explored these parameters on different sections and different types of movement for all infants. Once set, these parameters remained unchanged and were applied to all infants.

Fig. 4 shows a longer example of the PPG time series (Fig. 4A) with the estimated movement waveform (Fig. 4B). Additionally, it shows the binary markers of movement onset and offset of both the algorithm and the raters (see below) (Figs. 4C and 4D). As can be seen, there is a good match of the raw signal with the movement signal and the marking of movement onset and offset.

2.3.2. Bradycardia Correction—One of the challenges was that bradycardia episodes generated low-frequency fluctuations in the PPG signal as the slowing of heart rate rendered

periods >1.5s. As illustrated in Fig. 5, the mean pulse height during non-movement (Fig. 5A) and bradycardia (Fig. 5B) was higher than during movement (Fig. 5C). During bradycardia, the mean pulse width was also larger than that during non-movement segments. This information from the mean pulse height and width was used to first classify a segment as bradycardic. The Fourier period threshold for determining movement from the wavelet power spectrum of bradycardic segments was then increased from 1.5s to one that was linearly proportional to the width of the PPG cycles using min-max normalization. The new Fourier period threshold was obtained from the mean pulse width (mean heart rate) of the window as below:

$$\lambda_{new} = \frac{pulsewidth_{mean} - pulsewidth_{min}}{pulsewidth_{max} - pulsewidth_{min}}(\lambda_{max} - \lambda_{min}) + \lambda_{min} \quad (8)$$

Where $Pulsewidth_{mean}$ is the mean pulse width of the window; $Pulsewidth_{max}$, $pulsewidth_{min}$, λ_{max} and λ_{min} are scaling constants related to the bounds on pulse width and period threshold. $Pulsewidth_{max}$ and $pulsewidth_{min}$ were chosen as the pulse width corresponding to a mean heart rate of 125bpm and 150bpm, respectively. λ_{max} and λ_{min} were chosen as the period thresholds corresponding to 4s and 1.5s, respectively. This ensured that during any slowing down of the heart rate from 150 and 125bpm, the period threshold was scaled linearly from 1.5 to 4s. Any heart rate below 125bpm used a Fourier period threshold of 4s.

2.4. Manual Annotation by Raters

To validate this new approach of detecting infant movement, three raters annotated, movements from direct video observation along with viewing the time-series of the limb accelerometers and the mattress sensors. One of the raters is an author, a graduate student in Bioengineering who developed the annotation protocol. He trained the other two raters, who were undergraduate students in Behavioral Neuroscience at Northeastern University. All three raters performed the same procedure to retrospectively evaluate the recordings and were blinded to the results obtained from the wavelet analysis. The raters were not informed of each other's rating until after all rating was completed.

For the scoring, all acquired data were imported into a polysomnography software (Remlogic, Embla, Broomfield, CO) and the raters identified the onset and offset of movement. They also annotated them as either general gross-body movements or location-specific isolated limb movements. All raters followed the same annotation protocol: (1) The video recording was used to screen for the presence of movement. (2) For each epoch of movement appearing on the video, the time of movement onset was determined from the accelerometer and force sensor signals from the mattress; the onset time was noted in the worksheet. (3) The same procedure was used to detect the offset of movement. The onsets and offsets of each movement were subsequently exported for further analysis. In circumstances where a limb of the infant was occluded from the camera view, the accelerometer was used to identify movement. Intervals of interventions by clinicians, such

as feeding and medical procedures, had been annotated and excluded prior to the analysis. The total duration of data available for analysis from the 5 subjects was 20.6 hours.

The annotations from each rater were divided into 10s-epochs indicating the presence or absence of movement. The number of movement epochs was normalized to a total valid analysis time in hours. For example, a study with 4 hours (or 14400s) of valid analysis time was divided into 1440 10s-epochs; for each rater the number of epochs that included any movement was counted. This number was normalized per hour, since all infants had different recording times. The total number of 10s epochs in an hour is 360. The total movement duration scored by each rater as a percentage of total valid analysis time was also computed. The reliability of the quantification of the number of movement epochs and movement duration percent between the three raters was calculated using the intraclass correlation coefficient (ICC). The strength of the ICC was interpreted as: 0.00–0.25 negligible; 0.26–0.49 low; 0.50–0.69 moderate; 0.70–0.89 high; and 0.90–1.00 very high correlation.

The annotations from the three raters were combined into a single binary time-series showing the presence or absence of movement. The presence of movement required at least two raters to agree upon a time point as movement (Fig. 4). The dataset from all 18 infants, the annotations from the three raters, and the source code for the wavelet algorithm will be made available through www.Physionet.org

2.5. Performance Measures

The wavelet analysis and the raters' analysis rendered binary signals, indicating presence and absence of movement. Periods of interventions and feedings were excluded from both binary signals. Other exclusions were periods when the PPG probe was disconnected or rendered flat waveforms due to clipping of the PPG signal. A first step in the validation was to classify true and false detections on a point-by-point basis: the number of samples, where the values for both binary time series were 1: TP (true positive) and 0: TN (true negative); the number of samples, where the value for the algorithm was 0 and the raters' analysis was 1: FN (false negative) and vice versa: FP (false positive). Based on these numbers, sensitivity and specificity of the algorithm were computed as follows:

$$sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$Specificity = \frac{TN}{TN + FP} \quad (10)$$

Additionally, the ability of the wavelet-based method to detect the number of continuous movement segments was evaluated. A movement event was defined from the combined rater- based binary time-series showing a continuous segment containing 1's. The total number of movement events was represented by TM. The number of TM's that were detected by the wavelet algorithm was denoted by DM, whereas the number of false

detections was defined as FD. Based on these definitions, the true detection rate (TDR) and the positive predictive value (PPV) were calculated as:

$$TDR = \frac{DM}{TM} * 100 \quad (11)$$

$$PPV = \frac{DM}{(DM + FD)} * 100 \quad (12)$$

The raters had also annotated types of movements such as gross body movement or limb-specific movements (left/right and hand/leg). We used this information to further assess the TDR of our algorithm in detecting movements that originated from the limb connected to the PPG sensor versus isolated movements of limbs that were not connected to the PPG sensor. This analysis was performed with the individual annotations by each rater.

2.6 Application of the Algorithm to Study Maturation of Infant Movement Patterns

To determine whether the algorithm (validated in infants #1–5, Table 1) can be used to define movement features that correlate with infant development across PCA, we studied an additional group of infants (#6–18). Movement was classified into 3 categories depending on the duration of movement periods: <5s, 5–30s, and >30s. The percentage of movement time in each category was computed and normalized to the total analyzable duration. Periods of interventions and feeding were excluded prior to the analysis. Pearson's correlation analysis was used to examine the relationship between percentage of each movement category and PCA of infant on day of study. A p-value <0.05 was considered significant.

3. Results

3.1. Movement Annotation

From the five infants (#1–5), 20.6 hours of data were used to develop the movement algorithm, after periods of intervention and feed had been eliminated. Table 3 summarizes the annotation results obtained from the three raters, who were blinded to the movement signal as derived with the wavelet algorithm. The results of the wavelet algorithm were disclosed after all raters completed their scoring. It took each rater a minimum of 50 hours to complete all their annotations of the 5 infants. The intraclass correlation coefficient (ICC) calculated for the number of movement incidences showed a high agreement of reliability between the scorers with a value of 0.77 (C.I.: 0.06–0.97). However, the ICC for movement as a percent of total valid time was only 0.39 (C.I.: 0.19–0.89) indicating a low agreement between raters. The total percentage of movement time scored by Rater 3 was consistently higher than that scored by Raters 1 and 2. The reason for this discrepancy was that rater 3 combined separate movement events into one single movement event, which was based on appropriate subjective interpretation. To compensate for the inter-rater variability, inherent to subjective rating two raters to agree upon a time point as movement in keeping with previous studies³, indicating the presence or absence of movement. The presence of movement

required at least two raters to agree upon a time point as movement in keeping with previous studies.

3.2. Bradycardia Correction

Fig. 6 shows an example of the PPG time series during an epoch of bradycardia, together with the estimated movement waveform and the binary markers of movement before and after applying the correction. The PPG signal in Fig. 6A shows segments of non-movement, bradycardia and movement. The two time series show the movement signal (red) and the binary detection of 'movement' (black) during the bradycardia. In the two time series below (Fig. 6B), the bradycardia correction was applied and the segment with bradycardia was no longer indicated as movement by the binary marker. The lowest heart rate that could be detected using this method was limited by the width of the PPG window. This window was 1.6s in our algorithm, corresponding to a heart rate of 37.5 bpm which is much lower than the alarming levels in the NICU requiring nurse intervention.

3.3. Detection of Delay and Intra-Movement Gaps

The exemplary binary signals in Fig. 4 highlight two important additional elements of the time series: First, there was a small delay between the onset of the rater-based movement and the onset determined from the wavelet algorithm. This delay is again illustrated in Fig. 7A in a shorter segment of the accelerometer signal with the PPG, movement and binary signal, which showed a clear delay between the accelerometer and the PPG signal. This delay resulted from an instrument-related lag (~1.4s) in the PPG signal and a second lag related to the wavelet computation (~0.2s). This second feature arose from the fact that the algorithm tended to break a continuous movement into a sequence of separate shorter movements. This can be seen in Fig. 4C between the 10 to 20s time interval. The breaking of continuous movement into several segments may have occurred during slower or less forceful movements of the infant within a longer movement, where the PPG signal is not too distorted as seen in Fig. 4A. Therefore, the maximum power of the PPG segments during these sections was occasionally in the non-movement range of period. We therefore estimated the time for which the movement segments should be merged.

We estimated the delay and the artifactual intra-movement duration by optimizing the similarity between the wavelet-based and the combined rater-based binary time-series. The delay was varied between 0 and 3s in steps of 0.1s and the intra-movement merge duration was varied between 0 and 10s in steps of 0.1s. For every delay and intra-movement merge duration pair the values of TP, TN, FP, FN, sensitivity and specificity were computed. The optimal delay and intra-movement merge duration were selected as the ones that maximized the Youden index, calculated as $sensitivity + specificity - 1$. The Youden index was chosen since it gives equal importance to sensitivity and specificity. Any gaps between detected movements of duration less than the mean intra-movement duration from 5 subjects were merged.

Rather than using each subject's individual optimized parameters, the average estimated delay and intra-movement merge duration from all 5 subjects were used to compute all performance measures. This approach was chosen, rather than optimizing for each subject,

because we wished to apply the optimized algorithm to additional preterm infants (e.g., infants #6–18) from which movement information from video and sensors were not available. The estimate of delay was also subjectively confirmed by selecting a subset of movement events from the limb that held both the pulse oximeter sensor and an accelerometer. To ensure that a movement event was not an effect of the previous movement, the subset contained only those movements that were succeeded by at least 45s of non-movement period.

Fig. 7B shows the results for estimating the optimal onset delay and intra-movement merge duration for a single subject (#3). The Youden index for each variable combination is visualized by the heat map. For this subject, the maximum Youden index was achieved at a delay of 1.72s and with an intra-movement merge duration of 5.70s. The estimated delays and intra-movement merge durations for each subject are summarized in Table 4. The onset delay ranged between 1.60s and 1.84s across subjects with an average of 1.72s. The intra-movement merge interval ranged between 2.70s and 5.70s with an average of 3.86s.

3.4. Detection Performance

Using the average delay of 1.72s and average intra-movement merge duration of 3.86s, the sensitivity, specificity and Youden index for each subject were computed and listed in Table 4. The algorithm had a sensitivity of 79.6% (sd: 7.1%) and specificity of 81.1% (sd: 7.1%) in point-by-point comparison of movement with that of the combined rater-based detection. The performance of the proposed algorithm in detecting incidences of movement is presented in Table 5. In identifying incidences of movement, the algorithm had a TDR of 73% indicating that the algorithm could accurately detect 73% of the combined rater-annotated movement incidences. The PPV was 75.7%. The TDR of the algorithm in detecting incidences of movement that involved movement of the limb connected to the PPG and annotated by each other rater was 88.8% for rater 1 (Table 6), 74.1% for rater 2 and 82.4% for rater 3. Similarly, the TDR in detecting any isolated limb movement not connected to the PPG sensor and annotated by each other rater was 62.5% for rater 1 (Table 7), 52.8% for rater 2 and 52.1% for rater 3. This indicated that the algorithm was effective not only in detecting movements from the specific limb that it was attached to, but also from other limbs not directly connected to the PPG sensor.

3.5. Movement Development Across PCA

To investigate how the percentage of infant movement changed with PCA the Pearson's correlation was calculated for the data of 13 infants. Movements were parsed into three duration categories. Fig. 8 displays the results. Short bursts (<5s) showed a significant negative correlation with PCA ($r=-0.87$, $p<0.001$), indicating that movements such as twitches, startles and isolated short limb movements decreased as infants mature. While the intermediate movement durations (5–30s) showed no significant trend, the percentage of longer movements (>30s) significantly increased with PCA ($r=0.78$, $p<0.01$).

4. Discussion

Movement is an important sign of infant health and may provide clinicians with critical information on neurodevelopmental and cardio-pulmonary outcomes. Our work suggests that a routinely recorded signal, the photoplethysmogram, can serve as a continuous measurement of movement in preterm infants, revealing for the first time a method for monitoring a vital neurological function in this highly vulnerable population.

4.1. Optimization

We optimized the wavelet-based movement detection algorithm by empirically correcting for two sources of error: (1) delay in detecting onsets of movement; (2) gaps in movement signal caused by vigorous motion of the sensor.

The average movement onset delay was 1.72s. A small portion of this delay (~200ms) was due to a computational lag required to avoid the leading edge artifact of the wavelet transform. The remainder of the delay (~1.5s) was due to the processing of the PPG signal in the bedside oximeter.

Brief gaps in signal were another source of error, seen as a large, abrupt, and brief drop in PPG voltage during periods of vigorous movement. Our optimization procedure revealed that such gaps were due to movement only if the gaps were up to 3.86s. Longer intervals were more likely to reflect a true reduction in movement.

4.2. Previous Studies to Detect Movement

Several previous methods have attempted to detect motion artifact from PPG signals, with the purpose of improving estimation of heart rate and oxygen saturation. Poets and Stebbens developed an algorithm that compared heart rate (HR) measured from an electrocardiogram (ECG) with pulse rate calculated from PPG for short segments. Although artifact detection was reliable, the approach required an additional ECG channel. Others have used PPG signals to reduce motion and noise artifacts and improve the quality of PPG signals for accurate estimation of heart rate and SpO₂. Features in a support vector machine classification model have also been used to distinguish movement artifacts from cardiac pulsatile PPG signal. However, these techniques have not explored the usefulness of the signal to detect movement *per se* as a vital signal used for clinical assessment. Previously, we applied a wavelet method, based on our preliminary work, to movement detection in infants with Neonatal Abstinence Syndrome (NAS). The analysis did not account for the delay in detection and there were no errors in leading edge detection because only retrospective analysis was considered. Further, infants with NAS are not generally prone to bradycardia, hence effects of bradycardia were not addressed. The algorithm developed here can be implemented in real-time, corrects for any misclassified movement during bradycardia commonly seen in preterm infants. Further, it is optimized with respect to the delay in detecting the onset of movement and the intra- movement gaps that are inherent in the PPG method.

4.3. New Window to Study Development of Motor System

Our application of the movement detection algorithm to an additional group of premature infants (#6–18) found that the total duration of brief body movements (<5s) decreased with PCA, whereas the duration of longer body movements (>30s) increased. These changes indicate maturational changes in the motor control system where simple movements of short bursts, such as twitches, startles and sighs, progress into longer and more complex movements involving the entire body. Our findings are consistent with the developmental patterns found in other studies on premature infants that show a maturational reduction in brief movements”, and serve as a separate outcomes measurement that corroborates the validity of our algorithm in a clinical neurodevelopmental context.

4.4. Study Limitations

The algorithm had sensitivity and specificity of ~80% when compared with the rater-based detection of movement duration. While these are satisfactory values, there are important limitations inherent to rating movement using videography and movement sensors. Even though the video camera was positioned in the best possible location to capture the entire body of the infant, some extremities often moved outside the camera view or the field of view was obscured by objects used during routine care of the infant. Further, the accelerometer attached to the limbs was highly sensitive to abrupt movements, but insensitive to movements with constant velocity. The mattress-embedded sensors only detected gross body movements that distorted the mattress foam, but failed to detect other types of movement. Although subjective annotation can integrate movement from all sensors, they face an inherent limitation, of rater fatigue and other factors that reduce intra- and inter-observer agreement irrespective of the number of raters as reported in previous studies³. It is important to note that obtaining ground truth for human bodily movements is a fundamental and pervasive problem in all of human movement science. Even using additional markers such as 3D-motion capture of segmental joint movements is limited by occlusion and moving markers on the skin, and implementation of the method would face challenges in the NICU. Our approach in the current study was to use simpler subjective methods to optimize our algorithm. We then sought to test the validity of our algorithm by using a well-known clinical outcomes measure: changes in movement duration across post-conceptual age. The results of this application is consistent with the clinical literature”, and provides corroborating evidence that the optimized algorithm can yield clinically useful results.

5. Conclusions

Movement artifacts that degrade pulse photoplethysmography can be transformed to a signal that reliably measures movement activity in preterm infants. This wavelet-based movement detection algorithm is highly sensitive to gross body and limb movements and follows a maturational change that can be tracked for individual infants across a broad range of post-conceptual ages. Our findings suggest that a prospective study using PPG-based movement detection is feasible and can address important questions regarding the relationship between maturation of infant movement and development in a host of disorders, such as autistic and other cognitive disorders as well as a range of motor syndromes and diseases. Early

detection might lead to early interventions that can mitigate the severity of these devastating disorders. Furthermore, movement may also provide a vital index of cardio-pulmonary risks such as apnea, bradycardia and hypoxia, thereby enabling clinicians to implement early interventions that prevent such life-threatening events. Our study is a first step in this direction, introducing for the first time a readily accessible signal in the NICU. However, a larger prospective clinical study is needed to further optimize the algorithm, to better estimate its accuracy, and to predict neurodevelopmental outcome.

Acknowledgment

The authors thank Courtney Temple and Alan Gee for data collection; Adriell Louis and Hannah Taylor for data annotation; the NICU staff and physicians for subject recruitment; and James Niemi and his team at the Wyss Institute for constructing the movement sensor mattress. This work was supported by NSF SCH Grant #1664815, and NIH Grants R01-GM104987 and R21-HD089731, and the Wyss Institute at Harvard University,

References

1. Aarnoudse-Moens CSH, Weisglas-Kuperus N, van Goudoever JB, and Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics* 124:717–728, 2009. [PubMed: 19651588]
2. Barker SJ, and Shah NK. The effects of motion on the performance of pulse oximeters in volunteers. *Anesthesiology* 86:101–108, 1997. [PubMed: 9009945]
3. Bernhardt I, Marbacher M, Hilfiker R, and Radlinger L. Inter- and intra-observer agreement of Prechtl's method on the qualitative assessment of general movements in preterm, term and young infants. *Early Hum. Dev.* 87:633–639, 2011. [PubMed: 21616611]
4. Bickler PE, Feiner JR, and Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology* 102:715–719, 2005. [PubMed: 15791098]
5. Bruggink JLM, Van Braeckel KN, and Bos AF. The early motor repertoire of children born preterm is associated with intelligence at school age. *Pediatrics* 125:e1356–1363, 2010. [PubMed: 20457678]
6. Chan GSH, Fazalbhoy A, Birznies I, Macefield VG, Middleton PM, and Lovell NH. Spontaneous fluctuations in the peripheral photoplethysmographic waveform: roles of arterial pressure and muscle sympathetic nerve activity. *Am. J. Physiol. Heart Circ. Physiol* 302:H826–836, 2012. [PubMed: 22114133]
7. Chong JW, Dao DK, Salehizadeh SMA, McManus DD, Darling CE, Chon KH, and Mendelson Y. Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection-reduction approach. Part I: Motion and noise artifact detection. *Ann Biomed Eng* 42:2238–2250, 2014. [PubMed: 25092422]
8. Couceiro R, Carvalho P, Paiva RP, Henriques J, and Muehlsteff J. Detection of motion artifact patterns in photoplethysmographic signals based on time and period domain analysis. *Physiol Meas* 35:2369–2388, 2014. [PubMed: 25390186]
9. Fukumoto M, Mochizuki N, Takeishi M, Nomura Y, and Segawa M. Studies of body movements during night sleep in infancy. *Brain and Development* 3:37–43, 1981. [PubMed: 7258548]
10. Hadders-Algra M, Heineman KR, Bos AF, and Middelburg KJ. The assessment of minor neurological dysfunction in infancy using the Touwen Infant Neurological Examination: strengths and limitations. *Developmental Medicine & Child Neurology* 52:87–92, 2010. [PubMed: 19549207]
11. Hayes MJ, Plante LS, Fielding BA, Kumar SP, and Delivoria-Papadopoulos M. Functional analysis of spontaneous movements in preterm infants. *Dev Psychobiol* 27:271–287, 1994. [PubMed: 7926280]
12. Holditch-Davis D, Brandon DH, and Schwartz T. Development of behaviors in preterm infants: relation to sleeping and waking. *Nurs Res* 52:307–317, 2003. [PubMed: 14501545]

13. Krägeloh-Mann I, and Cans C. Cerebral palsy update. *Brain Dev* 31:537–544, 2009. [PubMed: 19386453]
14. Krishnan R, Natarajan BB, and Warren S. Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data. *IEEE Trans Biomed Eng* 57:1867–1876, 2010. [PubMed: 20172800]
15. Kuzniewicz MW, Wi S, Qian Y, Walsh EM, Armstrong MA, and Croen LA. Prevalence and neonatal factors associated with autism spectrum disorders in preterm infants. *J. Pediatr* 164:20–25, 2014. [PubMed: 24161222]
16. Mathew OP, Thoppil CK, and Belan M. Motor activity and apnea in preterm infants. Is there a causal relationship? *Am. Rev. Respir. Dis* 144:842–844, 1991. [PubMed: 1928959]
17. McGowan JE, Alderdice FA, Holmes VA, and Johnston L. Early childhood development of late-preterm infants: a systematic review. *Pediatrics* 127:1111–1124, 2011. [PubMed: 21624885]
18. Moortel ID, Munday SA, and Hood AW. Wavelet Analysis: the effect of varying basic wavelet parameters. *Solar Physics* 222:203–228, 2004.
19. <Munro BH Statistical methods for health care research Philadelphia: Lippincott Williams & Wilkins, 2005, xiii, 494 pp.
20. Poets CF, and Stebbens VA. Detection of movement artifact in recorded pulse oximeter saturation. *Eur. J. Pediatr* 156:808–811, 1997. [PubMed: 9365075]
21. Selvaraj N, Mendelson Y, Shelley KH, Silverman DG, and Chon KH. Statistical approach for the detection of motion/noise artifacts in Photoplethysmogram. *Conf Proc IEEE Eng Med Biol Soc* 2011:4972–4975, 2011. [PubMed: 22255454]
22. Shelley K and Shelley. S Pulse oximeter waveform: photoelectric plethysmography. *Clinical Monitoring*, Lake Carol Hines R, and Blitt C, Eds.: WB Saunders Company 420–428, 2001.
23. Shrout PE, and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428, 1979. [PubMed: 18839484]
24. Tobin RM, Pologe JA, and Batchelder PB. A characterization of motion affecting pulse oximetry in 350 patients. *Anesth. Analg* 94:S54–61, 2002. [PubMed: 11900040]
25. Torrence C, and Compo GP. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society* 79:61–78, 1998.
26. Williamson JR, Bliss DW, Browne DW, Indic P, Bloch-Salisbury E, and Paydarfar D. Individualized apnea prediction in preterm infants using cardio-respiratory and movement signals, 2013. doi:10.1109/BSN.2013.6575523
27. Zuzarte I, Indic P, Barton B, Paydarfar D, Bednarek F, and Bloch-Salisbury E. Vibrotactile stimulation: A non-pharmacological intervention for opioid-exposed newborns. *PLoS One* 12:, 2017.
28. Zuzarte I, Temple C, Indic P, and Paydarfar D. Transforming artifact to signal: A wavelet- based algorithm for quantifying neonatal movement. *Conf Proc IEEE Eng Med Biol Soc* 2014:5466–5469, 2014. [PubMed: 25571231]

Fig. 1.

Schematic illustration of force sensing resistor strips (blue) embedded in the mattress. The pulse oximeter was attached along with one accelerometer to a limb at the nurse's discretion. A second accelerometer was attached to another limb.

Fig. 2.

Exemplary raw data. A: Example of a PPG signal. The two segments S1 and S2 of 200 data points each (1.6s) mark a non-movement baseline (S1) and a period containing movement (S2). B: Estimated movement calculated with the wavelet-based algorithm. The vertical dashed lines point to the value of the estimated movement from each segment. Each of these is expanded in Fig. 3.

Fig. 3.

Graphical representation of the computation of movement. A and B: Example of PPG signals during non-movement (S1) and movement periods (S2) marked in Fig. 2. Both PPG segments were of length 200 samples (at a sampling rate of $125\text{Hz} = 1.6\text{s}$). C and D: Scalograms of the two segments using the Paul wavelet. Note that the period axis has a dyadic representation because period is a multiple of scale (Eqn. 6), which is expressed as powers of 2 (Eqn. 2). The y-axis of the scalogram represents Fourier period (or inverse of Fourier frequency). Higher values of period on the bottom left represent lower Fourier frequency values. The axis is reversed to depict low frequency values on the lower values of the axis. The scalogram gives a period distribution at each time point. Normalization of power was achieved using Eqn. 3 such that the wavelet function has unit energy at all periods. E and F: Time-averaged power spectrum from a 0.2 s time interval (vertical slice ranging from 3.6 to 3.8s in panel C and 9.6 to 9.8s in panel D). Since the power spectrum was averaged in time, the resulting average power is plotted as a function of period. The last 0.2s of the scalogram (between 3.8 to 4s in panel C and 9.8 to 10s in panel D) were excluded from the wavelet transform due to edge-effects. The dashed red lines indicate the period threshold of 1.5s. The maximum value of the average power at period $> 1.5\text{s}$ was the estimated movement for the selected PPG segments.

Fig. 4.

Wavelet-based and rater-based binary markers. A: Example of the PPG signal. B: Estimated movement from the wavelet-based algorithm. C: Binary markers of the wavelet algorithm determining the onset and offset of movements. D: Combined rater-based binary marker of movements. Three raters individually annotated movements by viewing video and time series of accelerometers and the mattress sensors. The combined rater-based binary signal was obtained retrospectively and required at least two raters to agree upon a movement at each time point. For the plotted time interval: Sensitivity = 82.2%, Specificity = 95.9 %, TM = 5, DM = 5, TDR = 100%.

Fig. 5.

Adaptation of the algorithm during bradycardia. A: PPG during the non-movement baseline segment. B: Bradycardia. C: Movement segment. D, E, F: Power extracted from the scalogram for each of the three conditions. Dashed lines in left and right panels indicate the period threshold at 1.5s for estimating movement; shifted to 4s proportional to pulse width during bradycardia.

Fig. 6.

Signals before and after correcting for bradycardia effects. A: PPG signal of 50s duration (blue) with bradycardia before 25s and fluctuations indicating movement after 25s. The bradycardia causes false detection of movement, as seen in the estimated movement (red) and the binary marker (black). B: Movement estimation with the bradycardia correction factor. The false detections in the first part of the signal (before 25s) no longer include movement markers.

Fig. 7.

A: Exemplary time series to illustrate the delay. Vertical dashed lines indicate the onset of movement from the accelerometer (left) and from the algorithm (right). Horizontal lines indicate the two components of delay: $d1$ from instrument related lag and $d2$ as the algorithmic delay. The total delay was estimated to be 1.6s. B: Optimization of the Youden index. The maximum Youden index was achieved at a delay of 1.52 s and with an intra-movement merge duration of 3.5s.

Fig. 8.

Movement duration as percent of total analyzable time over post-conceptional age (PCA) in three categories: <5s, 5–30s and >30s.

TABLE 1.

Subject Characteristics.

Subject ID	GA (wks)	PCA (wks)	Birth Weight (g)	Study Weight (g)	Gender	Study duration (hrs)	Analyzable duration (hrs)
1	31.57	32.85	1595	1490	M	4.39	2.73
2	26.71	30.85	1020	1390	F	5.16	3.28
3	27.28	29.00	880	906	M	8.56	4.40
4	31.28	32.00	1600	1550	M	8.17	5.30
5	27.85	29.42	1225	1240	F	8.05	4.92
6	27.29	29.43	1045	1220	M	45.62	39.60
7	30.57	30.71	1785	1755	M	43.84	36.38
8	30.14	30.71	1710	1710	M	43.71	37.86
9	29.43	30.14	925	843	F	46.78	41.03
10	31.29	32.29	1868	1674	M	49.75	36.91
11	28.86	30.14	1115	1140	F	48.61	39.91
12	28.57	30.14	1050	1110	M	20.34	17.29
13	32.14	32.43	2210	2100	F	24.60	21.03
14	30.29	30.57	1225	1230	F	70.32	53.43
15	33.86	34.29	1860	1900	M	47.27	36.74
16	28.71	30.29	1000	930	F	57.97	21.25
17	32.14	33.14	2030	1760	M	23.41	17.38
18	29.57	31.57	1295	1480	M	55.94	15.88

GA = Gestational age, PCA = Post-conceptual age

TABLE 2.**Criteria for Detection of Onset and Offset of Movement**

Criteria for movement onset: All features in condition 1		
	Standard deviation of preceding 25 sample (1s) window	> 3
Condition 1	Amplitude of current sample	> 10
	Slope of preceding 10 samples (0.4s)	> 40
Criteria for movement offset: Either condition 1 or condition 2		
Condition 1	Standard deviation of preceding 25 sample (1s) window	3
	Amplitude of current sample	10
Condition 2	Slope of preceding 10 samples (0.4s)	< -35

Thresholds are set given a PPG signal with minimum and maximum values of 1024 au and 3072 au respectively. For detecting the onset of movement, all three criteria under Condition 1 must be satisfied. For detecting the offset of movement, either Condition 1 or Condition 2 is sufficient.

TABLE 3.

Movement Percentage and Normalized Count Annotated by Raters

Subject ID	Movement (% of total valid time)			Number of 10-s movement epochs per hr		
	Rater 1	Rater 2	Rater 3	Rater 1	Rater 2	Rater 3
1	23.51	27.76	48.29	185.85	234.15	267.44
2	33.99	39.92	36.37	240.00	261.46	250.73
3	20.30	25.40	52.30	184.27	204.10	262.91
4	19.62	23.86	34.29	132.14	160.13	193.02
5	37.27	37.97	56.57	222.96	226.18	267.85

TABLE 4.

Similarity Measures

Subject ID	onset delay (s)	intra-movement merge duration (s)	Sensitivity *	Specificity *	Youden Index *
1	1.72	2.70	76.0	73.6	0.50
2	1.84	3.00	81.3	76.1	0.57
3	1.60	4.80	69.4	87.8	0.57
4	1.72	5.70	82.9	89.5	0.72
5	1.84	3.10	88.1	78.4	0.67
Average	1.74	3.86	79.6	81.1	0.61

* Measures computed using average onset delay of 1.74s and average intra-movement merge duration of 3.86s

TABLE 5.

Performance Measures for detection of movement events

Subject ID	TM	DM	FD	TDR (%)	PPV (%)
1	593	388	103	65.4	79.0
2	535	394	133	73.6	74.8
3	705	516	133	73.2	79.5
4	459	323	154	70.4	67.7
5	583	481	140	82.5	77.5
Average				73.0	75.7

TABLE 6.

Performance Measures (PPG-Limb Movements)

Subject ID	TM	DM	TDR (%)
1	297	223	75.08
2	430	368	85.58
3	494	451	91.30
4	334	314	94.01
5	501	458	91.42
Average			87.48

TABLE 7.

Performance Measures (Non-PPG-Limb Movements)

Subject ID	TM	DM	TDR (%)
1	144	96	66.67
2	85	36	42.35
3	106	64	60.38
4	56	36	64.29
5	80	60	75.00
Average			61.73