Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation



2012 Physiol. Meas. 33 1617

(http://iopscience.iop.org/0967-3334/33/10/1617)

IOP PUBLISHING PHYSIOLOGICAL MEASUREMENT

Physiol. Meas. **33** (2012) 1617–1629 doi:10.1088/0967-3334/33/10/1617

# Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation

**W Karlen[1], K Kobayashi[1], J M Ansermino[2] and G A Dumont[1]**

[1] Electrical & Computer Engineering in Medicine Group, Department of Electrical & Computer Engineering, The University of British Columbia (UBC), 2332 Main Mall, Vancouver, BC, V6T 1Z4, Canada.
[2] Electrical & Computer Engineering in Medicine Group, Department of Anesthesiology, Pharmacology & Therapeutics, The University of British Columbia (UBC), 2332 Main Mall, Vancouver, BC, V6T 1Z4, Canada.

E-mail: walter.karlen@ieee.org

Received 30 April 2012, in final form 1 August 2012
Published 18 September 2012
Online at stacks.iop.org/PM/33/1617

**Abstract**

Pulse oximeters are monitors that noninvasively measure heart rate and blood oxygen saturation ($SpO_2$). Unfortunately, pulse oximetry is prone to artifacts which negatively impact the accuracy of the measurement and can cause a significant number of false alarms. We have developed an algorithm to segment pulse oximetry signals into pulses and estimate the signal quality in real time. The algorithm iteratively calculates a signal quality index (SQI) ranging from 0 to 100. In the presence of artifacts and irregular signal morphology, the algorithm outputs a low SQI number. The pulse segmentation algorithm uses the derivative of the signal to find pulse slopes and an adaptive set of repeated Gaussian filters to select the correct slopes. Cross-correlation of consecutive pulse segments is used to estimate signal quality. Experimental results using two different benchmark data sets showed a good pulse detection rate with a sensitivity of 96.21% and a positive predictive value of 99.22%, which was equivalent to the available reference algorithm. The novel SQI algorithm was effective and produced significantly lower SQI values in the presence of artifacts compared to SQI values during clean signals. The SQI algorithm may help to guide untrained pulse oximeter users and also help in the design of advanced algorithms for generating smart alarms.

Keywords: photoplethysmography, signal quality index, segmentation, pulse oximeter, repeated Gaussian filters

(Some figures may appear in colour only in the online journal)

  

## 1. Introduction

Pulse oximetry allows for the noninvasive measurement of multiple vital signs with a single compact device. Heart rate (HR) is derived from the photoplethysmogram (PPG), a measure of arterial blood volume change , obtained by measuring the varying intensity of light traveling through the tissue. The distinct absorption properties of oxygenated and non-oxygenated hemoglobin at different wavelengths facilitate the noninvasive estimation of arterial blood oxygen saturation ($SpO_2$). Pulse oximetry has become a standard monitoring tool in many clinical settings.

Two significant limitations make reliable continuous patient monitoring with pulse oximetry challenging and can cause false alarms:

(i) the PPG is prone to motion artifacts and optical interference and
(ii) clear, reliable PPG signals are hard to obtain from patients with low perfusion (Murray and Foster 1996).

Clinical studies have shown that 71–77% of alarms from pulse oximeters are false positive alarms caused by patient movement (Petterson *et al* 2007). This high percentage of false alarms has resulted in higher workload for care givers and inefficient health care provision. While there has been extensive research on minimizing the effects of such artifacts in PPGs, these techniques fail to produce reliable values in the presence of significant artifacts and such failures must be detected (Lovell *et al* 2010). Methods to quantifiably measure perfusion exist, such as the perfusion index (Lima *et al* 2002); however, there are no methods for measuring artifacts.

We present the development of a real-time algorithm to calculate a signal quality index (SQI) for PPGs. The goal of an SQI is to provide an objective measure of the degree of signal corruption. This corruption could be due to artifacts or physiological states that decrease the reliability of the calculated vital signs. An SQI can be used as a real-time validation tool for high-level algorithms such as decision support or alarm engines. Alternatively, an SQI can be displayed on a user interface to provide additional information about the confidence of the displayed vital signs (HR and $SpO_2$). In addition, an SQI can be useful to those with no medical background, such as patients using a telehealth care application, to determine if they are using the pulse oximeter correctly.

### *1.1. Background*

While most algorithms described in the literature do not produce an SQI, they attempt to minimize the effects of artifacts in PPGs. Despite the difference in purpose, the approaches taken in the design of these algorithms can be applied in the design of an SQI algorithm.

Several algorithms segment PPGs according to heart beats before analyzing the signals. In such algorithms, the reliability of any further analysis is highly dependent on the accuracy of the beat segmentation (Rusch *et al* 1996, Krishnan *et al* 2008, Kim and Yoo 2006). Methods to segment PPGs using their derivatives have been discussed (Weng *et al* 2005, Farooq *et al* 2010). Furthermore, physiological limitations for maximum and minimum HR have been taken into account (interbeat-interval limitation) for the segmentation algorithm described in Farooq *et al* (2010); while this addition improves the beat detection rate, the method still has an error rate of 8.47%. Other approaches include multi-stage bandpass filters combined with nearest-neighbor decision logic and interbeat-interval limitation (Aboy *et al* 2005) or non-segmenting algorithms using adaptive frequency estimators (Karlen *et al* 2012) or Kalman filters (Seyedtabaii and Seyedtabaii 2008).

The use of waveform morphology of PPGs is one method of signal quality assessment presented in the literature (Sukor *et al* 2011). This method requires beat segmentation of the signals, then compares five attributes of each pulse: pulse amplitude, trough depth difference, pulse width, Euclidean distance with the average of previous pulses and amplitude ratios. By comparing these attributes to corresponding thresholds, pulses are given one of three quality ratings: good, poor and bad. This approach requires implementing a complex decision tree with multiple arbitrary thresholds. The selection of these thresholds can be challenging.

Principal component analysis (PCA) has been used for artifact removal (Hong Enríquez *et al* 2002) and extraction of HR and respiratory rate (RR) from PPGs (Madhav *et al* 2011). PCA is a statistical analysis tool that can be used to find patterns in a set of highly dimensional data, usually to compress the data by reducing the dimensionality. The kinds of data used for this analysis vary: some use PPGs segmented by beat, while others use lagged portions of the PPG. Independent component analysis (ICA) has been used to separate uncorrupted PPGs and artifacts from two signals obtained from the two LEDs in conventional pulse oximeter probes (Krishnan *et al* 2008, Kim and Yoo 2006). ICA is a variation of PCA which employs multiple mathematical models to separate independent signals from multiple signals that contain different mixtures of the independent signals.

Another approach is the analysis of PPG signals in the frequency domain. Numerous transforms exist that can output the power spectrum of signals. The power spectrum can be analyzed to obtain HR and $SpO_2$ and are very robust even in the presence of artifacts (Rusch *et al* 1996). Analysis in the frequency domain has also shown that clean data contain a lower number of strong frequency components compared to corrupt data (Krishnan *et al* 2008). Other signal sources such as electrocardiograms and accelerometers have been used to rate the quality of PPGs (Lovell *et al* 2010) or cancel out artifacts (Wood and Asada 2006, Lee *et al* 2010). However, this requires additional equipment or modification of the pulse oximeter probe.

## 2. Algorithm design

The algorithm was designed to be integrated into pulse oximeters used in clinical settings. This required the algorithm to be real time, and consequently must calculate the SQI iteratively and with as little delay as possible. The purpose of the algorithm is to rate the quality of the PPG. Therefore, the algorithm needs to be sensitive to artifacts, but not to other natural variations in the PPG such as heart rate variability (HRV), baseline variability (e.g. due to venous blood variations), or dicrotic notches (figure 1).

Our approach relies on beat segmentation of the PPG based on a mixture of time domain and frequency domain processing (figure 2). The SQI is then calculated for each pulse segment. In order to be quantitative, the SQI range has been selected to be between $-1$ and 100 in unit steps, where a value of $-1$ indicates that the algorithm is calibrating and an SQI is unavailable. A value of 0 corresponds to the worst signal quality, and a value of 100 corresponds to the best signal quality.

### *2.1. Beat segmentation*

The beat segmentation is based on *a priori* knowledge of the previous pulse rate and pulse amplitude. It is assumed that under non-artifact conditions each consecutive pulse is comparable and does not show much variation. Therefore the location of the next pulse can be estimated. This requires a robust method for estimating the initial HR.

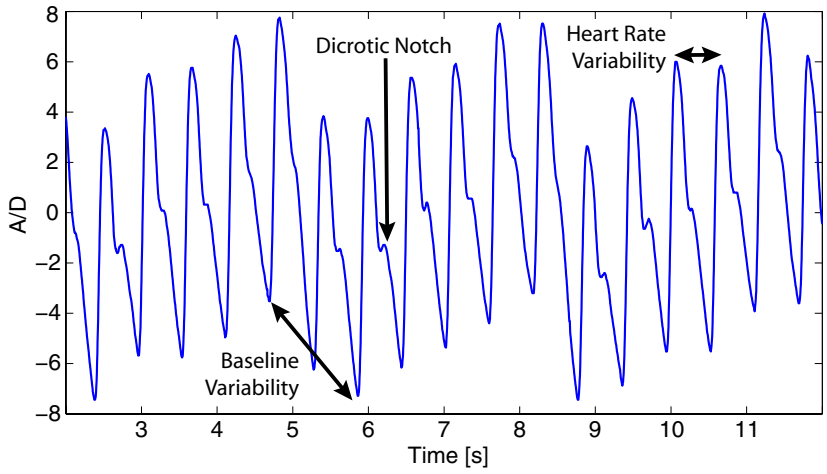


**Figure 1.** Photoplethysmogram with and without the presence of dicrotic notches. The baseline variation is present in this case.

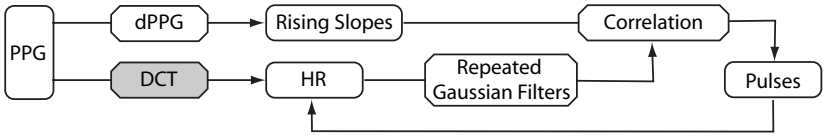


**Figure 2.** Pulse segmentation algorithm using repeated Gaussian filters. The photoplethysmogram (PPG) is processed using differentiation (dPPG) to identify rising slopes in the signal and the discrete cosine transform (DCT) to estimate the heart rate (HR). Correlation between the rising slopes and repeated Gaussian filter is calculated to identify the most likely position of the pulse. The DCT step displayed with a light gray background is computed if the HR is not available from previous iterations (calibration).

Initialization is completed using a calibration procedure. During the calibration of the beat segmentation, the HR is estimated using spectral analysis as it is more robust in the presence of artifacts than time-domain methods. The discrete cosine transform (DCT) was chosen to transform PPGs into the frequency domain because it does not require complex arithmetic and its frequency resolution is twice that of the discrete Fourier transform (Rusch *et al* 1996). The most dominant frequencies in the PPG are chosen to be estimates of the HR. Multiple estimates are evaluated to eliminate effects of large artifacts or other periodic components in the PPG signal. During calibration, the PPG is recorded until the pulse width estimation from DCT is three times smaller than the available PPG signal segment. This ensures that at least three pulses are available for processing. Calibration is also used after periods where no pulse was detected for longer than 6 s, such as after artifacts of longer duration or an asystole.

The derivative of the PPG (dPPG) is used to find the rising slopes in the PPG, which allows the identification of the primary pulse peak. We use a method that has already been described in Farooq *et al* (2010) and Weng *et al* (2005). The dPPG is obtained from

$$\mathrm{dPPG}_n = (\mathrm{PPG}_{n+1} - \mathrm{PPG}_{n-1}) * f_s/2, \tag{1}$$

where $f_s$ is the sampling rate. The dPPG output is always positive from the start of a pulse to the primary peak. However, a rising slope is not unique to the primary peak. It could also indicate the presence of secondary peaks caused by the dicrotic notch, high-frequency signal

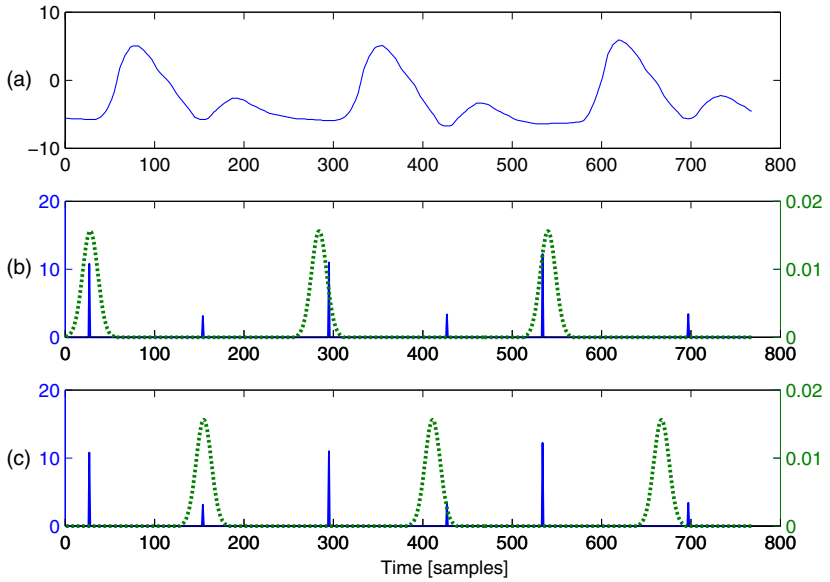


**Figure 3.** Example of the repeated Gaussian filter applied to a PPG signal with pronounced dicrotic notches. (a) PPG segment that was analyzed. (b) Test for the first possible offset by superposing the Gaussian filter (broken curve) on the marker for rising slopes (full curve) obtained from dPPG. (c) Test for the second offset. Offset from (b) is selected because correlation is higher.

components or artifacts. Therefore, *a priori* information about pulse rate and pulse amplitude is used to complete the beat segmentation.

Using estimates of HR obtained from DCT, repeated Gaussian filters are constructed. First, a single Gaussian is constructed such as

$$y = \frac{e^{-x^2}}{2 * \sigma^2}, \tag{2}$$

where $x$ ranges from $-3 * \sigma$ to $+3 * \sigma$, and where $\sigma$ is an adaptive fraction of the estimated pulse width. This Gaussian distribution is then duplicated and distributed at intervals of the HR estimates. An array from the dPPG is then constructed which indicates the height of each rising slope in the PPG. The repeated Gaussian filters are added with an offset for possible starting points of the first pulse, and a correlation score between rising slope array and a repeated Gaussian filter is obtained for each possible combination of HR and offset (figure 3). The combination that returns the highest correlation score determines the most likely HR and pulse offset.

When constructing the Gaussian filters there are two considerations that need to be made. First, the overall filter needs to be normalized so that if one of the estimated HRs was approximately twice the HR, the weight assigned to the secondary peaks is dampened by a larger decrease in the weight from the primary peaks. The risk of incorrectly estimating HR is thus decreased. Second, the overall filters should only contain integer multiples of Gaussian distributions. For example in figure 3, there are three whole distributions. If partial distributions were present, filters for the same estimated HR but different offsets may contain individual distributions that carry different weights. Having different weights for the distributions in each filter makes the comparison between the correlation scores with different offsets impossible.

If pulses were detected immediately before the portion of the PPG to be processed, the assumption is made that another pulse starts at the beginning of this portion and the previous

calculation of estimates for HR using DCT becomes unnecessary. Instead, the mean pulse amplitude and period of the previously detected pulses are used to construct the repeated Gaussian filters. This information is used to estimate the length of the next pulse and the height of the primary peak of the second pulse. All the rising slopes calculated from (1) are considered as candidates for the starting positions of the second pulse in the available data. Then, the location and the height of each rising slope are compared.

### *2.2. SQI calculation*

In continuous monitoring, a good signal quality means that the signal properties are stable over time and do not contain artifacts that could interfere with any post-processing method that changes the processing outcome. Two approaches can be taken for the calculation of an SQI; each pulse could be

(i) compared to previously detected pulses, or
(ii) rated individually using a static evaluator algorithm.

The first method is more favorable as it takes the shape of previous pulses into account and is adaptive. Superimposing pulses at their primary peaks reveals that, despite the differences in base amplitude and pulse widths, clean pulses taken from a short time frame are almost identical, except when corrupted with artifacts.

For calculating the SQI, we use the similarity in the pulse shapes within a time frame. Each new pulse is cross-correlated with the previous pulses at their corresponding primary peaks to obtain a normalized coefficient in the range of $-1$ to 1.

If a pulse is found to have a normalized cross-correlation coefficient ($C$) of more than the threshold th $= 0.99$, the pulse is determined as clean and added to the reference pulse set. The reference pulse set contains a maximum $n = 10$ last pulses. Larger sets are possible, but will increase processing load unnecessarily. When calculating the $C$ of a pulse, it is important to note that the pulse should always be longer than the reference pulses. Such a measure is necessary in order to consider the possibility that there are artifacts at the beginning or the end of the pulse being compared. In order to satisfy this condition, if a reference pulse is shorter than the pulse being compared, the length of the reference pulse needs to be increased. Weng *et al* (2005) have shown that the tails of PPG pulses can simply be extrapolated with the slopes of the tails. Since our beat segmentation method described above ensures that each detected pulse starts with the slope leading to the primary peak, the left tail can be ignored. Therefore when a reference pulse is shorter than the pulse being compared, only the right tail is extrapolated using the first-order linear regression.

The SQI is then obtained by applying a nonlinear scaling function to the $C$ of each detected pulse:

$$SQI = \frac{e^{\frac{50*(C+1)}{99}*8}}{e^8} * 100. \tag{3}$$

The factor 8 has been chosen to ensure that the SQI is well represented in the range of 0–100. An SQI of 0 corresponds to poor signal quality, while an SQI of 100 corresponds to the best signal quality where all pulses correlate. For computational efficiency, this has been implemented as a lookup table of size 101. Furthermore, if no pulse can be detected at any given time, as can occur with very irregular HRs or asystoles, an assumption is made that the signal quality of the data is too poor for any further analysis and an SQI of 0 is given for this duration. Although the SQI is meant to be in the range of 0–100, $-1$ is given specifically to indicate when the algorithm is calibrating at initialization, after a long sequence of artifacts, or after an asystole.

## 3. Methods

PPG data for testing and development were obtained from the Capnobase (CB) database (available at capnobase.org) maintained by the Electrical and Computer Engineering in Medicine group at the University of British Columbia (Karlen *et al* 2010) and from the 'pox' recordings in the Complex System Laboratory (CSL) database (available at bsp.pdx.edu) that were also used as a benchmark in Aboy *et al* (2005). The CB data were collected from subjects aged 1–74 that underwent general anesthesia. The signals were recorded using a Nellcor pulse oximeter sensor and sampled at 100 Hz using Datex–Ohmeda's S/5 Collect software. These signals have been upsampled to 300 Hz by the data export module. DC offsets have been removed in these signals. The CB database contains 124 cases that are 120 s in length and 42 cases that are 480 s in length. The 120 s long signals were assigned to a training data set that was used for algorithm development and calibration. The 480 s long signals from CB were used as a benchmark data set. For this study, a research assistant independently evaluated the beat annotations for their position and labeled all sections that contained artifacts. The CSL data set was used as second benchmark data set. It contained recordings obtained from two children in a pediatric intensive care unit and the beats were annotated by two different experts. All sections containing artifacts in the CSL data set were independently labeled by the same research assistant that labeled the CB data set.

Beat by beat comparison was performed using a pairing algorithm that follows the recommendations for Testing and Reporting Performance Results of Cardiac Rhythm by the American National Standard ANSI/AAMI (2003). Reference and algorithm peak locations were considered as matched if they were within a detection window of 30 ms. A learning period of 20 s was applied and the peaks within this period were not compared. The performance of the beat segmentation algorithm was assessed by measuring sensitivity (Se) and positive predictive value ($+P$) on the benchmark data sets such as

$$\mathrm{Se} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

$$+P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{5}$$

where TP is the number of true positives, FN the number of false negatives and FP the number of false positives. Se and $+P$ were compared to the output of the algorithm described in Aboy *et al* (2005).

To assess the quality of the SQI, normalized histograms of SQI for beats labeled by the expert as regular and artifact were plotted. The algorithms were implemented and all analysis performed using MathWorks' MATLAB version R2010b.

## 4. Results

A total of 27 992 (CB) and (CSL) pulses have been analyzed by the algorithm. 0.69% (CB) and 5.76% (CSL) of the beats were labeled by the expert as corrupt. While the CB benchmark contained less and shorter artifacts, this data set was more challenging to the algorithm because of its wider spectrum of subjects and pulse patterns.

### *4.1. Beat segmentation*

The beat segmentation showed similar performance for both benchmark data sets (tables 1 and 2). $+P$ was very high (>99.2%) and Se was above 96.2%. For the CB benchmark, only

**Table 1.** Confusion matrix for the CB benchmark set. $' > 50'$: percentage of pulses with an SQI>50. $' = -1'$; percentage of pulses during calibration SQI=−1.

| | | Expert annotation | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Pulse | | | No pulse | | | |
| | | count | $> 50$ (%) | $= -1$ (%) | count | $> 50$ (%) | $= -1$ (%) | +P (%) |
| Algorithm | Pulse | 25 688 | 96.1 | 2.7 | 51 | 21.6 | 5.9 | 99.80 |
| | No pulse | 948 | 4.3 | 92.0 | – | | | |
| | Se (%) | 96.44 | | | | | | |

**Table 2.** Confusion matrix for the CSL benchmark set. $' > 50'$: percentage of pulses with an SQI>50. $' = -1'$; percentage of pulses during calibration SQI=−1.

| | | Expert annotation | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Pulse | | | No pulse | | | |
| | | count | $> 50$ (%) | $= -1$ (%) | count | $> 50$ (%) | $= -1$ (%) | +P (%) |
| Algorithm | Pulse | 15 170 | 94.6 | 4.8 | 120 | 29.2 | 50.8 | 99.22 |
| | No pulse | 597 | 8.2 | 80.9 | – | | | |
| | Se (%) | 96.21 | | | | | | |

**Table 3.** Confusion matrix for the CSL reference algorithm from the CSL benchmark data set.

| | | Expert annotation | | |
|---|---|---|---|---|
| | | Pulse | No pulse | +P (%) |
| Algorithm | Pulse | 15 613 | 336 | 97.89 |
| | No pulse | 160 | | – |
| | Se (%) | 98.99 | | |

4.3% of the missed pulses had an SQI>50. The majority of the missed pulses were produced during calibration (92%). A similar pattern was found for the CSL benchmark.

The CSL benchmark data set provided the output of the CSL reference algorithm (table 3). The presented algorithm showed higher $+P$ and a lower Se than the reference algorithm.

### *4.2. SQI*

An example of the SQI algorithm's output is shown in figure 4. The figure shows an example of the calibration stage, and also a decreased SQI due to an artifact. The SQI obtained for pulses with the artifacts are clearly lower than the SQI of other pulses.

The SQI at the regions labeled as artifacts by an expert and regions not labeled as artifacts were compared to rate the validity of the SQI algorithm (figure 5). Over 90% of the regions labeled as artifacts are given an SQI lower than 95, whereas only 14% of the regions not labeled as artifacts are given an SQI less than 85. The results also show that, in the presence of artifacts, SQI is significantly lower, resulting in an SQI = 0 in the majority (54%) of cases.

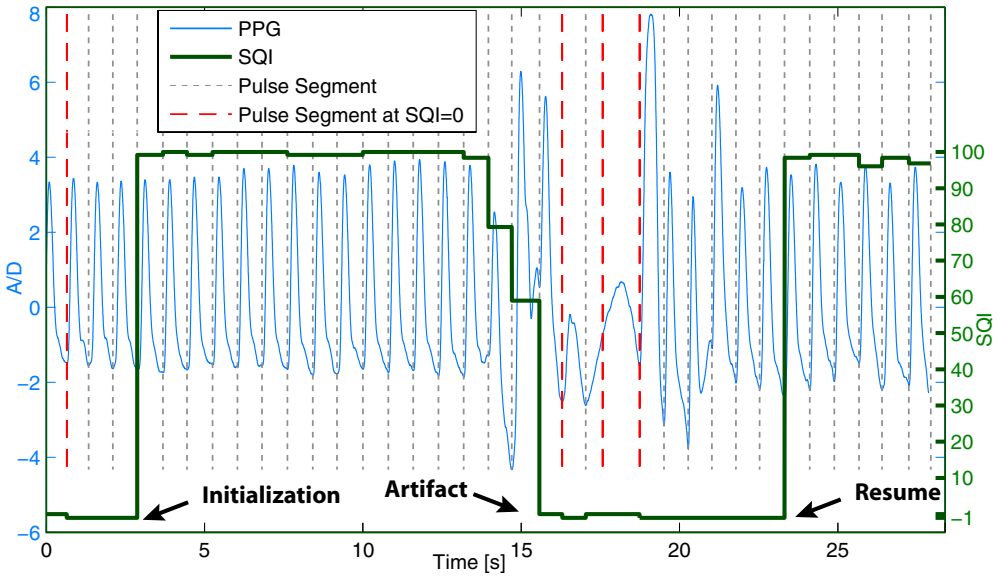


**Figure 4.** Example of signal quality index (SQI) calculation at the beginning of a case followed by an artifact in the photoplethysmogram (PPG).

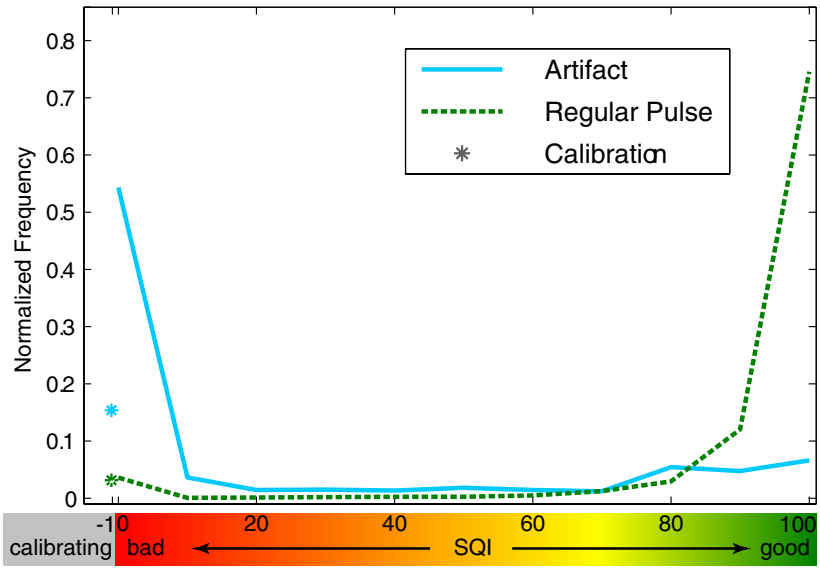


**Figure 5.** Distribution of SQI for pulses labeled as artifact and regular on the combined CB and CSL benchmark data sets. The color coding of the SQI is for illustration purposes only.

### *4.3. Delay*

During calibration, the beat segmentation algorithm requires a PPG signal with at least three times the length of the estimated pulse. In addition, the SQI algorithm requires at least one reliable reference pulse for calculating the cross-correlation. Thus, at any time, the SQI cannot

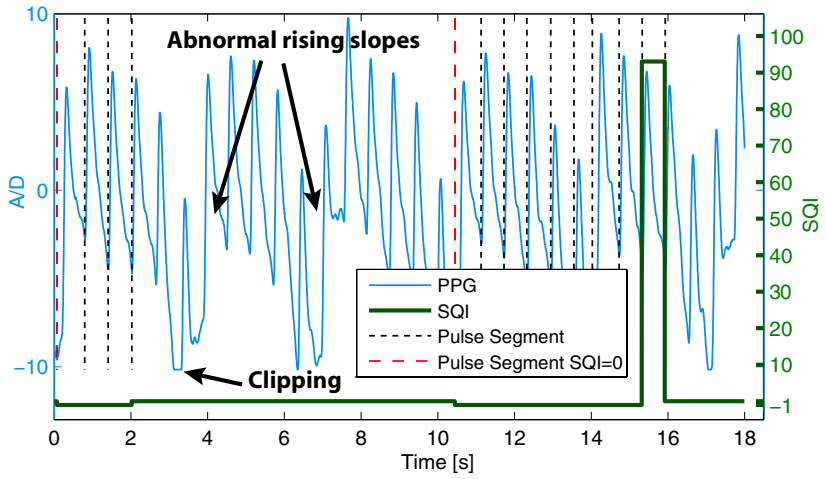


**Figure 6.** Sequence of the lowest performing case of the CB benchmark set. Many pulses in the PPG are not detected because of large baseline variations that resulted in clipping and abnormal pulse amplitudes.

be obtained for the past two or more pulses with the number of pulses depending on the HR estimated by the beat segmentation algorithm. This delay of minimum three pulses can be seen in figure 4 at initialization or after the artifact. In difficult cases as shown in figure 6, the delay was up to seven pulses.

## 5. Discussion

We have described a novel approach to determine an SQI in a PPG signal. Existing signal segmenting approaches have been combined with a novel method using repeated Gaussian filters for localizing segments of pulses. Using this pulse segmentation, an algorithm that calculates an SQI using cross-correlation for each detected pulse was presented.

The algorithm produced an SQI that correlated well with the occurrence of artifacts in the PPG. During the presence of artifacts, the SQI was significantly lower compared to the SQI during a clean PPG signal. This makes the SQI calculation method very effective in identifying artifacts and providing information of reliability of the PPG signal. The presented SQI algorithm is not intended to be only for artifact removal. It can be used independently or in conjunction with artifact removal algorithms in order to rate the PPGs and any statistics of features obtained from signals. It is of high value to subsequent signal processing and decision support algorithms used for clinical applications. They could use the SQI information to improve output and condition it to specific conditions. For example, alarm procedures can adjust priorities of the generated alarms based on the SQI and consequently reduce false alarms during low signal fidelity. An SQI that is sensitive to artifacts and other signal disturbances could also be valuable for user interfaces where direct feedback about signal quality can be given to the user. The SQI algorithm does not use a discrete threshold to discriminate between poor and bad PPG pulses. It would rather be the application that defines ‘good enough’ and ‘too poor’ signals. Therefore, we have not provided a definition of a single threshold in this manuscript and used a generic limit that was at the center of the SQI range (SQI = 50) for the presentation of results. Once an application has been defined, our data could be used to

establish a threshold. For example, a histogram such as displayed in figure 5 could allow for the selection of a threshold and with the support of receiver operating curves, the optimal sensitivity and specificity could be determined. Some applications may also demand more classes than only 'good' and 'poor'.

As mentioned in the algorithm design section, a requirement was that the SQI output was not to be affected by HRV. Repeated Gaussian filters satisfied this criterion by considering not only intervals of estimated HRs but regions around the estimated starting locations of pulses (figure 3). In the case of arrhythmias, a new pulse may not be found at the expected location despite the Gaussian filter. For example, in the case of asystoles, extrasystoles and sudden onsets of bradycardias the algorithm will not segment the PPG signal and the SQI will decrease to 0. In the case of sudden onset of tachycardias, the segments will likely contain multiple pulses. As a result, the cross-correlation with the previous reference pulses will return a very low score and the SQI will also decrease. We believe that a lower SQI during these conditions is appropriate as they are uncommon and need attention. For example, the calculated HR is normally an average over multiple beats and will be affected by arrhythmia. A lower SQI would highlight this.

The novel pulse segmentation algorithm showed lower Se than $+P$. The decrease in Se can be attributed to the nature of the algorithm. In the regions with significant artifacts, where the repeated Gaussian filter would not match with the detected peaks, the algorithm does not segment the signal, even if multiple clear peaks are visible. The experts' visually inspected signals contain annotated peaks also during artifact periods, although it might arguably be better to ignore these peaks for further processing. It is therefore preferable to optimize the algorithm for high $+P$, instead of high Se. Furthermore, after a larger artifact, the algorithm requires a few pulses to lock-in to the new HR which can cause a delay in the segmentation. This more selective automated peak annotation produced a more robust solution to artifacts with three times less FP. Considering that only a few FN occurred during normal SQI $>50$ (8.2%), a low SQI is a good indicator of the fact that pulses were missed by the algorithm.

A limitation of the segmentation algorithm was its prolonged difficulty of locking-in and finding the correct HR in one CB case. This happened for the following two reasons:

(i) the baseline variability was significant compared to the AC component of the PPG and caused abnormally large amplitudes in the rising slopes at respiratory inspiration (figure 6), very likely due to mechanical control ventilation; and

(ii) the large variability in baseline contributed to lower weight carried by each rising slope in the multiplication with the repeated Gaussian filters due to each rising slope falling at the edge or outside each Gaussian distribution.

In the example shown in figure 6, both of these conditions were largely present and thus, the beat segmentation failed over a longer period of time. In addition, clipping was present in the signal that produced a low correlation coefficient. During this period, the SQI remained low and prevented a false HR calculation due to wrong segmentation.

We will continue to work on future improvements of the SQI algorithm that take these limitations into account. One possible way to decrease the risk of a beat segmentation failure is to widen each Gaussian distribution and/or make each distribution more fat-tailed. Thus, a large-scale sensitivity analysis comparing the beat segmentation accuracy with different parameters will be required.

The presented results were from data obtained from the operating room and the intensive care unit where subjects were anesthetized or sedated. Consequently, motion artifacts were less frequent than in awake subjects. A total of 1094 corrupted beats were available for

comparison. Further validation on awake subjects with longer and more pronounced motion will be required.

The pulse segmentation using the novel repeated Gaussian filter algorithm was compared well with other existing algorithms. The $+P$ was higher and the Se was slightly lower than the CSL reference algorithm presented in Aboy *et al* (2005). However, the reported numbers are difficult to be compared as the CSL algorithm was not designed to evaluate the quality of the signal or detect artifacts.

Further work will consist of reducing the computational cost of the algorithm. While the currently presented algorithm was implemented in Matlab and performed well in real time, further optimization is necessary for an embedded or mobile device. One possibility of decreasing the processing time is to decrease the number of PPG pulses to use as reference beats for the beat segmentation algorithm. By decreasing this threshold, the maximum amount of data that needs to be processed in one iteration is directly decreased. This will help reduce the time lag of the SQI that can occur in some instances.

## 6. Conclusion

The calculation of a signal quality index (SQI) for PPG is an important task and can be performed independently of artifact removal. Low SQI can indicate poor algorithm performance caused by various reasons and increased likelihood for false positives and false negatives. This is valuable information for subsequent signal processing algorithms and users alike. The presented novel pulse segmentation and SQI algorithm based on repeated Gaussian filters and cross-correlations are a promising approach to provide information from pulse oximetry more reliably.

## Acknowledgments

The authors would like to express their gratitude to Erin Cooke who provided the expert annotation to the benchmark data and Joanne Lim for her helpful feedback on various versions of this manuscript. This work has been supported by the Swiss National Science Foundation (SNSF), Grand Challenges Canada and the Canadian Institutes of Health Research (CIHR).

## References

Aboy M, McNames J, Thong T, Tsunami D, Ellenby M S and Goldstein B 2005 An automatic beat detection algorithm for pressure signals *IEEE Trans. Biomed. Eng.* **52** 1662–70

ANSI/AAMI 2003 Testing and reporting performance results of cardiac rhythm and ST-segment measurement algorithms *ANSI/AAMI EC57* (Arlington: Association for the Advancement of Medical Instrumentation)

Farooq U, Jang D G, Park J H and Park S H 2010 PPG delineator for real-time ubiquitous applications *Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (31 Aug–4 Sept 2010, Buenos Aires, Argentina)* pp 4582–5

Hong Enríquez R, Sautié Castellanos M, Falcón Rodríguez J and Hernández Cáceres J L 2002 Analysis of the photoplethysmographic signal by means of the decomposition in principal components *Physiol. Meas.* **23** N17–29

Karlen W, Petersen C, Gow J, Ansermino J M and Dumont G A 2012 Photoplethysmogram processing using an adaptive single frequency phase vocoder algorithm *BIOSTEC 2011: 4th Int. Joint Conf. on Biomedical Engineering Systems and Technologies (Communications in Computer and Information Science* vol 273*)* ed A Fred, J Filipe and H Gamboa (Berlin: Springer) pp 31–42

Karlen W, Turner M, Cooke E, Dumont G A and Ansermino J M 2010 CapnoBase: Signal database and tools to collect, share and annotate respiratory signals *STA: Annu. Meet. of the Society for Technology in Anesthesia (West Palm Beach)* p 25

Kim B S and Yoo S K 2006 Motion artifact reduction in photoplethysmography using independent component analysis *IEEE Trans. Biomed. Eng.* **53** 2001–3

Krishnan R, Natarajan B and Warren S 2008 Motion artifact reduction in photopleythysmography using magnitude-based frequency domain independent component analysis *Proc. 17th Int. Conf. on Computer Communications and Networks ICCCN'08 (3–7 Aug 2008, St. Thomas, US Virgin Islands)* pp 1–5

Lee B, Han J, Baek H J, Shin J H, Park K S and Yi W J 2010 Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry *Physiol. Meas.* **31** 1585–603

Lima A P, Beelen P and Bakker J 2002 Use of a peripheral perfusion index derived from the pulse oximetry signal as a noninvasive indicator of perfusion *Crit. Care. Med.* **30** 1210–3

Lovell N H, Redmond S J, Basilakis J and Celler B G 2010 Biosignal quality detection: an essential feature for unsupervised telehealth applications *12th IEEE Int. Conf. on e-Health Networking Applications and Services (Healthcom)* (1–3 July 2010, Lyon, France) pp 81–5

Madhav K V, Ram M R, Krishna E H, Reddy K N and Reddy K A 2011 Extraction of respiratory activity from PPG and BP signals using principal component analysis *Int. Conf. on Communications and Signal Processing (ICCSP 2011) (10–12 Feb 2011, Kerala, India)* pp 452–6

Murray W B and Foster P A 1996 The peripheral pulse wave: information overlooked *J. Clin. Monit.* **12** 365–77

Petterson M T, Begnoche V L and Graybeal J M 2007 The effect of motion on pulse oximetry and its clinical significance *Anesth. Analg.* **105** (Suppl. 6) S78–84

Rusch T, Sankar R and Scharf J 1996 Signal processing methods for pulse oximetry *Comput. Biol. Med.* **26** 143–59

Seyedtabaii S and Seyedtabaii L 2008 Kalman filter based adaptive reduction of motion artifact from photoplethysmographic signal *World Acad. Sci. Eng. Technol.* **37** 173–6

Sukor J A, Redmond S J and Lovell N H 2011 Signal quality measures for pulse oximetry through waveform morphology analysis *Physiol. Meas.* **32** 369–84

Weng J, Ye Z and Weng J 2005 An improved pre-processing approach for photoplethysmographic signal *27th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (1–4 Sept 2005, Shanghai, China)* pp 41–4

Wood L B and Asada H H 2006 Noise cancellation model validation for reduced motion artifact wearable PPG sensors using MEMS accelerometers *28th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (30 Aug–3 Sept 2006, New York, NY, USA)* pp 3525–8