A Pattern-Based Windowed Impulse Rejection Filter for Nonpathological HRV Artifacts Correction

Hussein Al Osman, Mohamad Eid, and Abdulmotaleb El Saddik, *Fellow, IEEE*

Abstract—Artifacts in a heart rate variability (HRV) signal can severely distort the extracted time- and frequency-domain parameters, and thus render the information obtained from the signal potentially unusable. In this paper, we propose an algorithm for nonpathological HRV artifact detection called pattern-based windowed impulse rejection (PWIR) filter. This algorithm extends the windowed impulse rejection filter algorithm, we have introduced in our previous publication. Our performance evaluation demonstrates that PWIR compares favorably with respect to the existing algorithms with a true detection rate of 93.80% and a false detection rate of 0.66%. Integral pulse frequency modulation, the most prominent competing algorithm according to our tests, displays a lower true detection rate (87.62%) and a significantly higher false positive detection rate (4.25%). All other algorithms are shown to be significantly less accurate. We also propose an interpolation technique to replace erroneous samples with new ones. The interpolation technique is shown to produce more accurate sample estimates than several existing ones (in terms of maximizing the accuracy of corrected HRV signals when it comes to time and frequency-domain analysis). The combination of the proposed artifact detection and interpolation methods presented in this paper constitutes a nonpathological artifact correction solution for HRV signals.

Index Terms—Biological signal processing, electrocardiography (ECG), heart rate variability (HRV), signal filtering.

I. INTRODUCTION

THE measurement of the heart's electrical activity, formally called electrocardiography (ECG), has been used extensively for the past century in the monitoring and diagnosis of cardiovascular ailments [1]. From the plethora of information that can be retrieved from an ECG signal, we are particularly interested in heart rate variability (HRV). HRV refers to changes in the time intervals between successive heart beats (or QRS complexes) [2], [3]. The usefulness of such signals was initially demonstrated in [4], where they showed that a decrease in HRV was associated with fetal distress.

Ever since, HRV has gained a reputation for somewhat indicating the general health of the heart [5], but also importantly, for reflecting important aspects of the autonomic processes [6]. In practice, the analysis of HRV has been mostly used to assess early prognosis of cardiovascular diseases [5], [7] and measurement of mental workload and stress [8], among others.

According to the Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology report [9], several types of parameters, based on the time or frequency domains may be computed through HRV analysis. Time-domain analysis includes the calculation of the mean value, the standard deviation (SDNN), the standard deviation of averaged values in segments (SDANN), and the average of standard deviations in segments. The frequency domain analysis includes the calculation of the power spectral density of the HRV signal that is divided into several bands: Ultra Low Frequency (0.0–0.0033 Hz), Very Low Frequency (0.0033–0.04 Hz), low frequency (LF) (0.04–0.15 Hz), and high frequency (HF) (0.15–0.4 Hz), as well as the ratio of power in LF band to that in HF band (usually called the LF/HF balance).

Commonly, a short-term HRV measure consists of 3 to 5 min worth of ECG data. Numerous mobile ambulatory ECG sensors can be used for that purpose (such as the ones described in [10]–[12]). The ECG data are then processed using an R-peak detector algorithm to identify all the QRS complexes. Many algorithms have been developed for that purpose [3], [13]–[15]. The series of R to R peaks intervals then constitute a signal we refer to as RR time series. The variations in the latter series are known as HRV [9]. This time series can be directly used to calculate time-domain parameters, however, it needs to be sampled at periodic intervals before being converted to the frequency domain to extract frequency-domain parameters.

Short HRV measurements are ideal for biofeedback applications that relay instant information regarding the autonomic processes to the user. In particular, applications that monitor instantaneous mental stress make use of such measures [16], [17]. Nonetheless, some HRV sequences contain artifacts, such as ectopic beats, arrhythmic events, noise, and other sources of artifacts that severely deteriorate the accuracy of the time- and frequency-domain analysis of the signal [18]. Various methods have been proposed to detect artifacts, such as the ones described in [19]–[22], [24]–[26], [31], and [32].

However, these algorithms still have their limitations under specific circumstances as it will be shown in this paper. In addition, some methods have been proposed to replace erroneous points with estimated ones to reduce the effect of these artifacts on time- and frequency-domain analysis. These are referred to in this paper as interpolation techniques. McNames *et al.* [19] and Lippman *et al.* [23] provide examples of such methods.

The remainder of this paper is organized as follows. Section II presents the sources of nonpathological artifacts and reviews related algorithms for HRV artifacts detection and interpolation. In Section III, we describe five existing artifact detection schemes and four existing interpolation methods. We also present our proposed method for correcting artifact plagued HRV signals that relies primarily on artifact detection and interpolation. Section IV comprises the evaluation of the presented an HRV detection and interpolation approaches, in comparison with existing ones. Finally, Section V summarizes the paper's findings and provides perspectives for future works.

II. BACKGROUND AND RELATED WORKS

Many types of cardiovascular arrhythmias can severely affect the integrity of an HRV signal. If the effect of such arrhythmias is not isolated, very little useful information can be yielded from the analysis of the signal. This paper focuses only on HRV signals collected from subjects that do not exhibit a history of heart disease. Usually, the purpose of HRV data collected from such subjects is mostly to assess fatigue, mental stress level, and wellbeing of the autonomic nervous system (ANS) in general. If we discount pathological arrhythmias, then an artifact in an HRV signal can mostly originate from two probable sources as follows.

1) *Ectopic Heartbeats:* These are disturbances in the cardiac rhythm that occur when the heart's electrical activity does not originate in the sinoatrial node.
2) *Measurement Noise:* HRV signals are extracted from ECG sequences. Any degradation in the quality of the ECG signal or measurement interruption can be reflected as an artifact in the HRV record. Moreover, any ECG feature incorrectly identified as a QRS complex by the R-peak detection algorithm also produces an HRV artifact.

A. Related Artifacts Detection Techniques

Several artifact detection algorithms have been proposed. For instance, Lee and Yu [22] and Houtveen and Molenaar [26] suggest the use of a simple thresholding (ST) technique, where each HRV sample is compared with predefined and static lower and upper bounds. Any sample falling outside of this bound is considered erroneous. Nonetheless, such approach has a limited efficiency on dynamic signals, where the heart rate frequently changes. Many ectopic beats for instance produce RR intervals that fall well within the bounds of the normal heart rate range. (This became particularly apparent from our analysis of such method in Section IV.)

Another approach is proposed in [20] and [21] and makes use of a sliding average window filter. In [21], for each sample,

the window consisting of the previous 30 s worth of data is averaged to produce a value upon which a lower and upper thresholds are based. The sample is considered to be an artifact if it falls outside the boundary of these thresholds. In [20], the algorithm involves the selection of a window size of $(2N + 1)$ data points, averaging the N points on either side of the central point, excluding the central point if it lies a specific fraction outside of window average, then advancing to the next data point. Algorithms [20] and [21] are fairly similar, therefore, we have chosen to only further investigate algorithm [21] in Section IV of this paper.

Since all non-pathological HRV artifacts are manifested as impulses, McNames *et al.* [19] propose an impulse rejection (IR) filter to identify erroneous samples. Similarly, Al Osman *et al.* [25] present the windowed impulse rejection (WIR) filter that segments the signal into overlapping windows and applies an IR filter on each one of them. The motivation for such a procedure is to produce an IR filter that adjusts to the dynamic nature of the heart rate by analyzing small sections of the signal at a time, instead of examining it as a whole.

Alternatively, Mateo and Laguna [24] base the artifact detection method on the integral pulse frequency modulation (IPFM) model. The IPFM model has been suggested to simulate the work of the ANS in modulating the heart activity, and thus the algorithm in [24] exploits this property to detect ectopic beats.

Furthermore, the work presented in [31] describes an online artifact detector that buffers limited amounts of samples, while checking for erroneous ones. Finally, Berntson *et al.* [32] present an artifact detector that employs thresholds principally dependant on an estimate of the standard deviation. The latter estimate is less corruptible by artifacts (compared with the regular standard deviation).

Sapoznikov *et al.* [33] present methods to reduce the effects of various pathologic arrhythmias on HRV analysis. Nonetheless, our work focuses on nonpathologic sources of HRV artifacts, and therefore the latter methods will not be considered.

B. Related Interpolation Techniques

Interpolation techniques are means to replace artifacts with new data points that are ideally equal in value to the samples that would have registered had the artifacts not occurred. Unfortunately, in the real-world some information is missed after correction. Nonetheless, estimating these samples by applying various mathematical and heuristic methods can yield good results, especially that sudden heart rate changes do not occur instantaneously (at least according to the IPFM model [24]).

Linear and cubic interpolation techniques are widely used to replace unwanted samples in numerous types of applications. Nonetheless, in terms of interpolation techniques specifically targeted for HRV artifact replacement, McNames *et al.* [19] propose a simple median based method. It involves the replacement of an artifact with the median of a window surrounding it. Conversely, Lippman *et al.* [23] use a predictive interpolation technique for this purpose. The predictive technique involves
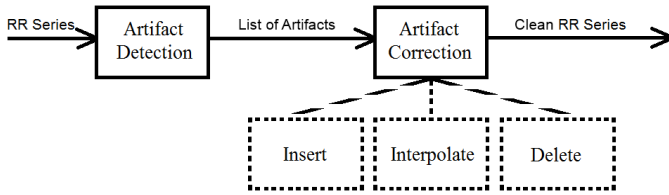
Fig. 1.   RR time series artifact filtration process.

the search for an artifact free sequence of data points that resembles the one plagued with the artifact. Once a relevant sequence is found, the erroneous data point is replaced by a sample from that sequence.

C. Conclusion

In this paper, we will investigate the effectiveness of all the aforementioned techniques in terms of their accuracy at detecting HRV artifacts or interpolating samples to replace unwanted ones. In regards to artifact detection methods, in addition to investigating the percentage of true detection, we will also examine their rate of false positive identification. False positives refer to the erroneous identification of a correct data point as an artifact. Moreover, we will present a new method for artifact detection (PWIR) and analyze how it fares compared with the existing ones. We will also present our interpolation technique (cubic predictive interpolation) and compare it to common and existing ones. Therefore, by proposing the latter two methods, we therefore provide a complete framework for correcting nonpathological artifacts in an HRV signal.

III. METHODS

A. Artifact Filtration

Canceling the effects of artifacts on time- and frequency-domain HRV analysis is performed in two stages. In the first stage, the artifact detection, a list of all the artifacts in an RR time series is compiled. In the second stage, the artifact correction, the effect of the artifacts is rectified using one of three operations: insertion, deletion, and interpolation of a sample (Fig. 1). Details regarding the context of the latter operations will be provided later in Section III-G.

B. Categories of Artifacts

To better understand the causes of artifacts, whether they have a physiological or measurement source, the following three sections divide these artifacts into three categories and provide an overview of their origins. The categories are formed according to how the artifacts are manifested in the RR time series, rather than their source. For instance, an interruption in the measurement or a QRS miss by the R-peak detection algorithm produce a similar structure, and therefore these artifacts belong to the same category.

In this paper, we focus solely on the types of artifacts that are commonly generated though nonpathological sources. Therefore, the categories of artifacts that will be discussed

in the following sections do not summarize all the possible structures of artifacts that can be found in HRV records.

However, before we introduce the three categories of artifacts, we must define the following parameters:

$$I_1 = X(a)$$
$$I_2 = X(a+1)$$
$$I_3 = X(a) + X(a+1)$$
$$\bar{X} = \frac{\sum_{i=1}^{N} X(i)}{N}$$

where $X(i)$ is the RR time series (sequence of all RR intervals), a is the index of the artifact and N is the number of samples in the RR time series.

1) Category 1: This category of artifacts is manifested in the RR time series as an upward peaking impulse composed of a single data point. This can be caused by an interruption of measurement or a miss of one or more QRS complexes by the R-peak detection algorithm (commonly referred to as false negative). In the latter case, an interpolation is required to estimate the location of the undetected or unmeasured R-peak(s). Achieving such outcome by manipulating the RR time series consists of removing the data point causing the impulse and replacing it with two or more interpolated data points. Fig. 2 displays a category 1 artifact, along with its manifestation in the ECG and RR time series. Note that the dashed circle indicates the presence of the artifact.

2) Category 2: This category of artifacts is manifested as a downward peaking impulse composed of two or more data points [Fig. 3(b)]. The cause of such structure might be an ectopic beat mistaken for a normal one or an ECG feature (such as, a T wave) incorrectly identified as an R peak (commonly referred to as false positive). In practice, from our examination of such artifact structure, we have found that it generally is composed of two points.

In the case of an ectopic beat, the extra contraction causes the heart timing to restart, and therefore, after a short delay firing again from the sinoatrial (SA) node. This produces strangely spaced normal beats with an ectopic beat in between [Fig. 3(a)]. The interval between the two normal beats (I_3) is noticeably smaller than two times the average interval between consecutive normal R-peaks ($2 \times \bar{X}$). On the other hand, I_3 is in most cases noticeably larger than $\bar{X}$. This produces a downward peaking impulse because the interval between the first normal beat and the ectopic beat (I_1) is noticeably smaller than $\bar{X}$. Likewise, the interval between the ectopic beat and the second normal beat (I_2) is also noticeably smaller than $\bar{X}$. In the case of an incorrectly detected R-peak between normal R peaks, I_3 is usually in the vicinity of $\bar{X}$.

3) Category 3: These artifacts exhibit a downward peaking impulse composed of one data point followed by an upward peaking impulse composed of one data point as well [Fig. 4(b)]. Such structure is almost exclusively caused by an ectopic beat. In this case, the ectopic beat prevents the normal one from taking place, and therefore the heart skips a normal beat. The next normal beat resumes the regular schedule of the heart electrical activity and comes right, where it would have been had the ectopic episode not occurred [Fig. 4(a)].
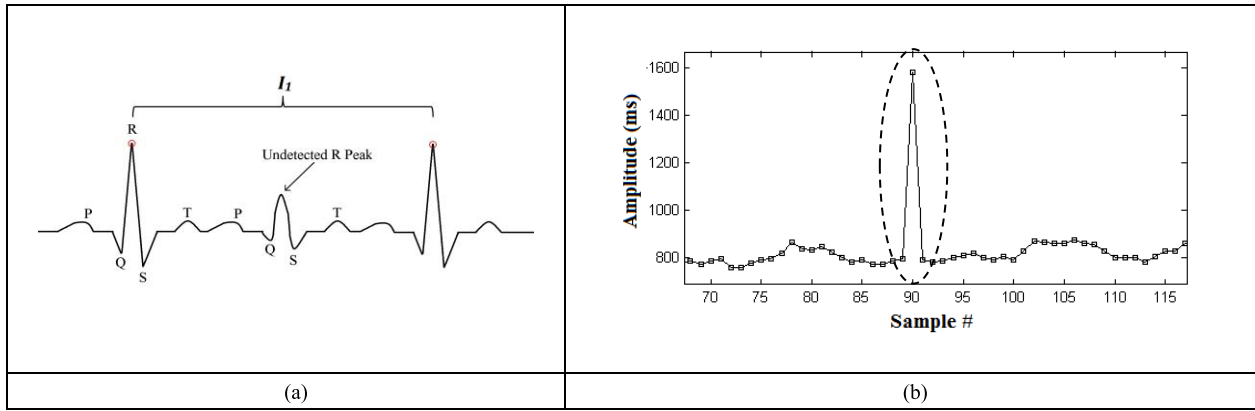
Fig. 2. Category 1 artifact in the ECG signal and the RR time series. (a) ECG manifestation. (b) RR time series manifestation.
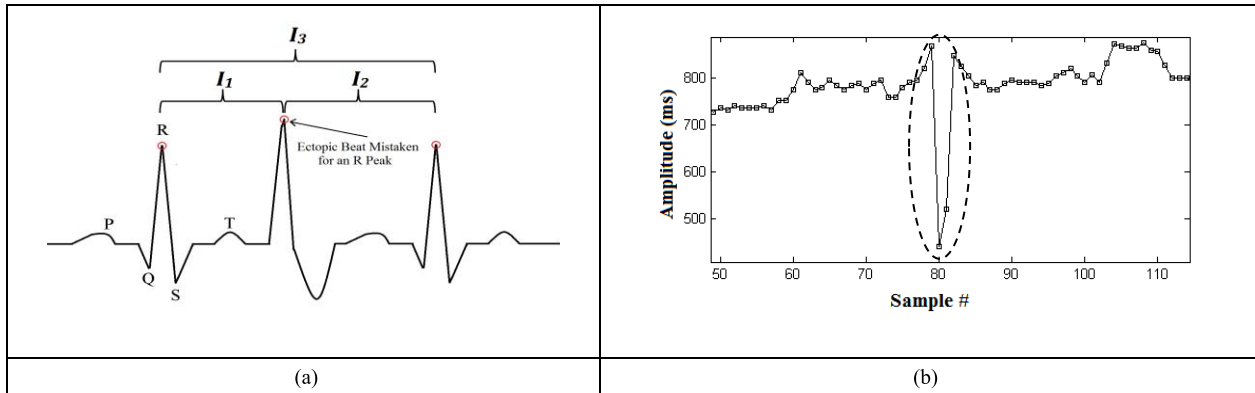

Fig. 3. Category 2 artifact in the ECG signal and the RR time series. (a) ECG manifestation. (b) RR time series manifestation.
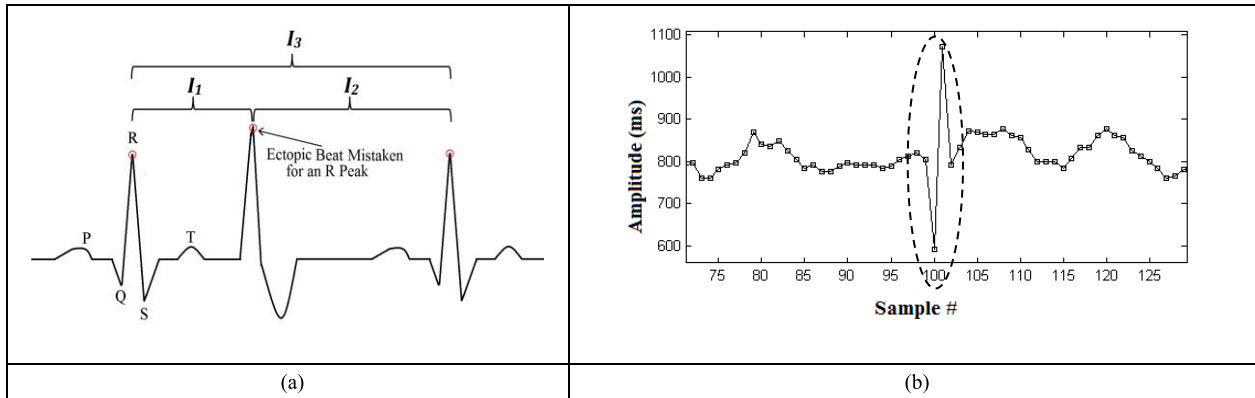

Fig. 4. Category 3 artifact in the ECG signal and the RR time series. (a) ECG manifestation. (b) RR time series manifestation.

In this case, the interval between the two normal beats surrounding the ectopic beat is almost two times the average of all RR intervals ($I_3 \approx 2 \times \bar{X}$). This results in a downward peaking impulse caused by the unusually small interval between the normal beat and the ectopic one (I_1), followed by an upward peaking impulse caused by the unusually long interval between the ectopic beat and the next normal beat (I_2). Therefore, in this case, the interpolation algorithm must replace the two erroneous data points with two interpolated ones.

C. Existing Artifact Detection Algorithms

In this section, we will present five existing algorithms for detecting artifacts in an HRV signals. In the following section,

we will introduce our own proposed method that extends the WIR filter presented in [25].

1) Simple Thresholding (ST) Algorithm: Perhaps the most basic algorithm to identify artifacts in a list of RR intervals makes use of an upper and lower threshold [22], [26]. The idea is that normal heart rate is limited in range, and therefore, detecting RR intervals that violate such paradigm indicates the presence of an artifact. This is formalized using

$$S(i) = \begin{cases} X(i), & R_1 \le X(i) \le R_2 \\ \text{Correct}\{X(i)\}, & \text{otherwise} \end{cases} \quad (1)$$

where $S(i)$ is the resulting filtered RR time series, R_1 and R_2 are the upper and lower thresholds, respectively.
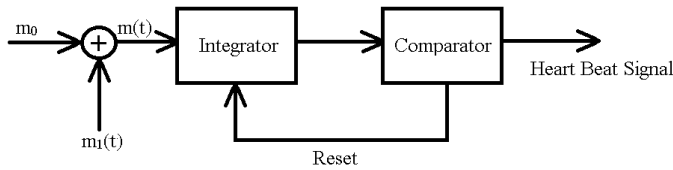
Fig. 5. IPFM model [28].

Houtveen and Molenaar [26] use a value of 300 ms for R_1 and 1800 ms for R_2.

2) Moving Average (MA) Algorithm: This algorithm is based on the work presented in [21]. For each sample, a window containing the previous 30 s of HRV data points is averaged to produce an expected interbeat interval (EIBI) value for each RR interval. The EIBI series is therefore calculated using (2), where n represents the previous 30 s worth of RR interval samples (i.e., $\sum_{j=i-n}^{i} X(j) = 30000$ ms)

$$\text{EIBI}(i) = \frac{\sum_{j=i-n}^{i} X(j)}{n}. \tag{2}$$

Then, an erroneous sample is detected and corrected using the following test:

$$s(i) = \begin{cases} X(i), & \theta \times \text{EIBI}(i) \leq X(i) \leq \alpha \times \text{EIBI}(i) \\ \text{Correct}\{X(i)\}, & \text{otherwise} \end{cases} \tag{3}$$

where $\alpha \in\ ]1, \infty[$ and $\theta \in\ ]0, 1[$. Mulder [21] sets α to 2 and θ to 0.6. The idea behind this algorithm is simple; each new sample must be somewhat similar to a window of samples that preceded it. If a sample is therefore found to be significantly different than its predecessors, it is judged to be erroneous, and therefore corrected.

On the other hand, a special arrangement has to be made for the first 30 s of the sequence, as there are not enough samples that precede them to fill a window. For these particular samples, the EIBI value produced from the first window in the record can be used in the calculation of their thresholds.

3) Integral Pulse Frequency Modulation (IPFM) Algorithm: This algorithm is presented in [24] and uses the IPFM model to detect ectopic beats. IPFM has been considered to describe the behavior of the ANS in controlling the heart rate [24]. At the root of this model is the modulating signal $m(t)$ that replicates the sympathetic and parasympathetic influences on the firing of the SA node [27]. Therefore, a heartbeat is generated when the integral of (4) reaches a threshold Ψ

$$\Psi = \int_{t_i}^{t_{i+1}} m(t)dt \tag{4}$$

where t_i is the time of the previous beat and t_{i+1} is the time of the new beat. The modulating function $m(t)$ is composed of a dc component m_0 and an dynamic component $m_1(t)$. Similarly, such phenomenon can be represented by the block diagram of Fig. 5 [28].

Now, Mateo and Laguna [24] exploit the fact that the variation in the instantaneous heart rate due to nonectopic beats is band limited according to the IPFM model.

Therefore, using a thresholding test on the derivative of the instantaneous heart rate, ectopic beats can be detected (as ectopic beats cause a faster change in the instantaneous heart rate than normal ones). The instantaneous heart rate derivative is estimated using

$$I_k^{'} = 2 \left| \frac{t_{k-1} - 2t_k + t_{k+1}}{(t_{k-1} - t_k)(t_{k-1} - t_{k+1})(t_k - t_{k+1})} \right| < U\sigma \tag{5}$$

where $I_k^{'}$ represents the derivative of the instantaneous heart rate, t_k is the time of the kth heart beat, U is a predefined threshold, and σ is the standard deviation of the RR time series. Consequently, a beat is considered nonectopic when the derivative of the instantaneous heart rate ($I_k^{'}$) is bigger than the threshold ($U\sigma$). Mateo and Laguna [24] set the value of U at 4.3.

4) Impulse Rejection (IR) Filter Algorithm: This algorithm is presented in [19]. It makes use of an IR filter that utilizes the median function to detect any possible artifacts. The following test statistic is employed:

$$d(i) = \frac{|X(i) - Xm|}{1.483 \times \text{med}\{|X(i) - Xm|\}} \tag{6}$$

where med $\{.\}$ is the median operator applied over the entire record of $|X(i) - Xm|$ and Xm is the median of $X(i)$. The filtered signal is then calculated as

$$S(i) = \begin{cases} X(i), & d(i) < \phi \\ \text{Correct}\{X(i)\}, & d(i) \geq \phi \end{cases} \tag{7}$$

where $S(i)$ is the resulting filtered RR intervals sequence and ϕ is a predefined threshold. McNames *et al.* [19] set the value of ϕ to 4.

5) Windowed Impulse Rejection (WIR) Filter Algorithm: One of the problems in the previous algorithm (IR) is that it is not sensitive to major changes in the heart rate. It looks at the record as a whole. Many ECG sensors, such as [10]–[12], are wearable, and therefore monitor the heart's behavior, while the user performs various activities. An example of such application is an ECG monitor that an athlete might wear during exercising to collect her or his heart rate and HRV parameters. During such activities, the heart rate tends to change considerably. Nonetheless, such drastic changes in the heart rate in a short time span are not accounted for in the previous algorithm as we see in Section IV.

The aforementioned weakness can be overcome using a windowed approach that segments the record into small portions for evaluation, as presented in [25]. McNames *et al.* [19] propose windowing for long records (with each window containing at least 5 min worth of samples). Nonetheless, Al Osman *et al.* [25] propose the use of small overlapping windows to account for the potential dynamic nature of the HRV signal. The motivation for overlapping the windows is to take into regard the signals continuous nature [25]. It would be disadvantageous to treat adjacent windows completely independently, while in reality one is a continuation of the other. Using (8) and (9), the number of windows needed to
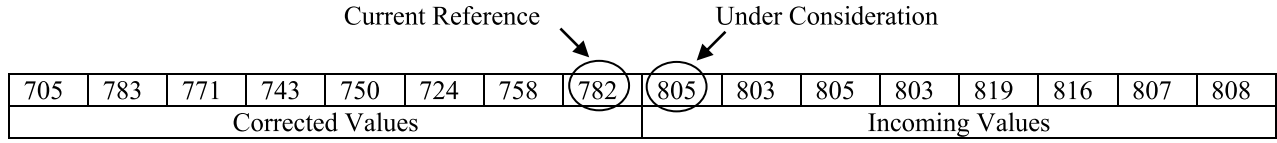
Current Reference   Under Consideration

705	783	771	743	750	724	758	782	805	803	805	803	819	816	807	808
Corrected Values								Incoming Values							

Fig. 6.   Snapshot of the RTC engine's buffer [31].

filter the signal is calculated

$$d = \lceil \omega \times (1 - \alpha) \rceil \qquad (8)$$

$$l = \left\lceil \frac{(N - \omega)}{d} \right\rceil + 1 \qquad (9)$$

where α is the overlap factor that has the range $0.5 \leq \alpha < 1$ (for an overlap of at least 50%), ω is the window length, N is the total number of samples in the signal, and l is the number of windows. This guarantees that every sample is tested at least twice with regards to its surroundings to judge whether it is an artifact (except for the first d samples that are tested only once). We define $W_j(h)$ as the content of the jth window with the first element corresponding to $X(j \times d)$ and the last element corresponding to $X((j \times d) + \omega)$.

For each window, the following array is calculated:

$$P_j(h) = \frac{|W_j(h) - Wm_j|}{\text{med}\{|W_j(h) - Wm_j|\}} \qquad (10)$$

where $1 \leq h \leq \omega$, med $\{.\}$ is the median operator applied over the entire record of $|W_j(h) - Wm_i|$ and Wm_j is the median of $W_j(h)$. The filtered signal is therefore calculated using the following:

$$S((j \times d) + h) = \begin{cases} W_j(h), & P_j(h) < T \\ \text{Correct}\{W_j(h)\}, & P_j(h) \geq T \end{cases} \qquad (11)$$

where $S(i)$ is the resulting filtered RR time series. In [25], α was set to 0.5, ω to 30 and T to 7.

6) Real-Time Corrector (RTC) of Interbeat Interval Algorithm: Rand *et al.* [31] introduce an engine for the online correction of RR intervals that uses an adaptive threshold. The engine maintains a buffer that stores the latest RR intervals measured (Fig. 6). The size of the buffer can be modified by the operator. The buffer is divided into two parts.

1) *Corrected Values Part:* Contains the values that have already been corrected.
2) *Incoming Values Part:* Contains the values that have been received by the engine but not corrected yet.

The absolute difference between the RR interval under consideration and current reference is verified against an adaptive threshold. Both of these values are labeled in Fig. 6. Once the value under consideration is dealt with, a left shift is performed on the buffer to load more value(s) onto it. The threshold is recalculated for each new RR interval under consideration using

$$AT = \sum_{i=0}^{NC-2} \frac{|b(i) - b(i+1)|}{NC - 1} m. \qquad (12)$$

Equation (12), AT is the adaptive threshold, NC is the number maintained in the corrected values part of the buffer,

$b(i)$ represents the values in the same part of the buffer as well, and m is a multiplier that can be set by the operator. In addition to generating an adaptive threshold, the engine makes sure that the threshold falls between two fixed bounds referred to as min(AT) and max(AT). In [31], values 5, 10, 50, and 200, where proposed for NC, m, min(AT), and max(AT), respectively.

7) Berntson's Algorithm: Berntson *et al.* [32] attempt to find a statistical limit (SL) for the largest acceptable absolute difference between consecutive RR intervals using statistical methods. First, the standard deviation of the RR intervals series is estimated using

$$SD = 1.48QD \qquad (13)$$

where SD is the estimated standard deviation and QD is the quartile deviation. Note that the standard deviation is estimated rather than computed since artifacts tend to skew the true standard deviation of the series; nonetheless, the provided estimation is less sensitive to artifacts. Second, the maximum expected difference (MED) and the minimal artifact difference (MAD) are calculated using

$$MED = 3.32QD \qquad (14)$$

$$MAD = \frac{Xm - 2.9 \times QD}{3}. \qquad (15)$$

Equation (15), Xm is the median of the RR intervals series. Finally, the SL against which the absolute difference between consecutive values of the series will be judged is calculated as the average between MED and MAD. Additional measures detailed in [32] are taken to reduce the FPDR.

D. Proposed Artifact Detection Algorithm

In this paper, we propose the PWIR filter algorithm for artifact detection in an HRV signal. This technique uses the same impulse detection mechanism as the WIR algorithm; nonetheless, in addition, we also examine the patterns formed by sequences of samples to assess whether they match any of the artifact categories discussed in Section III-B. Therefore, the procedure works as follows.

1) Segment the vector $X(i)$ into overlapping windows using the same method introduced by the WIR algorithm.
2) Calculate a Normalized Vector for each window using

$$NW_j(h) = \frac{W_j(h) - Wm_j}{\text{med}\{|W_j(h) - Wm_j|\}}. \qquad (16)$$

Note that (16) is very similar to (10) except the sign of each value in the produced vector is preserved; this is imperative to allow later for the recognition the

TABLE I
TESTS TO ASSESS THE CATEGORY OF THE POTENTIAL ARTIFACT

Category Test	Potential Category
$m_1 > 0$ AND $m_2 < 0$	1
$m_1 < 0$ AND $m_3 > 0$	2
$m_1 < 0$ AND $m_2 > 0$ AND $m_3 < 0$	3

TABLE II
CUSTOMIZED ARTIFACT TESTS FOR EACH CATEGORY

Category	Condition for Correction
1	$|m_1| \geq T_1$ AND $|m_2| \geq T_1$
2	$|m_1| \geq T_2$ AND $|m_3| \geq T_2$
3	$|m_2| \geq T_3$ AND $(|m_1| \geq T_3$ OR $|m_2| \geq T_3)$

categories of potential artifact. In plain language, the normalized vector divides the difference between each sample in $W_j(h)$ and the median of $W_j(h)$ by the median of all such differences.

3) Consider every sample $NW_j(h)$ to correspond to an artifact in $W_j(h)$ (until proven otherwise) and test for its potential category. Note that $W_j(h)$ is the jth window in $X(i)$ and corresponds to values between $X(j \times d)$ and $X((j \times d) + \omega)$ (d is calculated using (8) and ω is the window size). The latter can be achieved through the tests described in Table I, where m_k (for $0 \leq k \leq 2$) is calculated using

$$m_k = NW_j(h + k) - NW_j(h - 1 + k). \quad (17)$$

Therefore, by looking at the slopes associated with every four consecutive points, we can discriminate which potential category of artifact they form.

4) If the sample, along with its surroundings, forms a structure resembling one of the artifact categories, then perform the corresponding test in Table II to verify if the sample is an erroneous one, and therefore should be corrected.

In Table II, the T_1, T_2, and T_3 thresholds are positive. By checking for patterns and impulses, rather than just impulses in a window, we can set much stricter threshold values without worrying about increasing the rate of false detections (i.e., reaching the false conclusion that a valid sample is an artifact). Hence, we are not only looking for sudden jumps in the HRV signal but also at the shape formed by every consecutive four samples (Fig. 7).

E. Existing Interpolation Techniques

In this section, we present four existing interpolation techniques. In the subsequent section, we present our proposed interpolation technique. Note that interpolation is the principal corrective action that is undertaken during the artifact correction phase (Fig. 1).
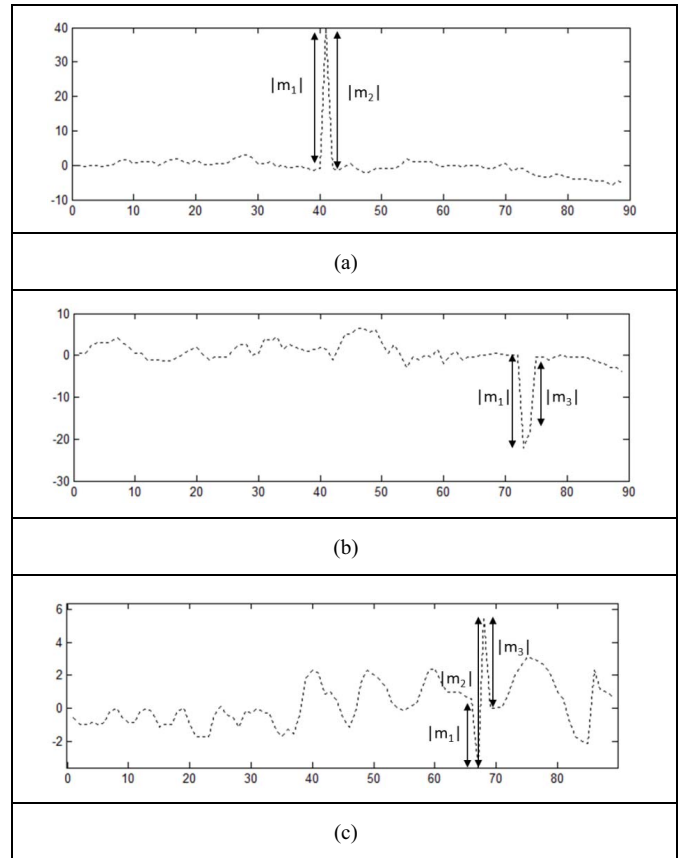

(a)

(b)

(c)

Fig. 7. Magnitudes of m_1, m_2, and m_3 for normalized vectors containing artifacts. (a) Normalized vector showing a category 1 artifact. (b) Normalized vector showing a category 2 artifact. (c) Normalized vector showing a category 3 artifact.

1) Median Interpolation: A mechanism for interpolating samples to replace artifacts is proposed in [19] and uses

$$\hat{X}(a) = \text{med}\{W_m(i)\} \quad (18)$$

where $\hat{X}(a)$ is the newly interpolated value, $W_m(i)$ is a window that corresponds to the samples between $X(a - \omega - 1/2)$ and $X(a + \omega - 1/2)$, and ω is the length of the window. McNames *et al.* [19] used a window of size $5\omega = 5$.

2) Linear Interpolation: Perhaps the simplest method to perform interpolation is through a linear function. Some small portions of the HRV signal (composed of few contiguous data points) display a somewhat linear behavior. This is especially true for HRV signals collected during normal or fast breathing. Slow breathing rates tend to increase the LF and decrease HF components [29] and consequently creating a less linear, more curvy, signal. However, this method provides simplicity in implementation and calculation.

McNames *et al.* [19] provide an interpolation technique based on the median of the window surrounding the detected artifact. The interpolated sample is calculated using

$$\hat{X}(a) = \frac{X(a - 1) + X(a + c)}{2} \quad (19)$$

where $\hat{X}(a)$ is the newly interpolated value, a is the index of the sample to be interpolated, and c is the number of consecutive artifacts starting at $X(a)$.

3) Cardinal Cubic Spline Interpolation: The following method is based on the cardinal splines interpolation method, which in turn is based on the more generic hermite interpolation. First, the four hermite basis functions are calculated using the following:

$$h_1(s) = 2s^3 - 3s^2 + 1 \tag{20}$$
$$h_2(s) = -2s^3 + 3s^2 \tag{21}$$
$$h_3(s) = s^3 - 2s^2 + s \tag{22}$$
$$h_4(s) = s^3 - s^2. \tag{23}$$

The value of s is calculated as follows:

$$s = \frac{i_c}{c + 1} \tag{24}$$

where c is the total number of consecutive points to be interpolated starting at $X(a)$ and i_c is the index in the latter string of artifacts. Therefore, $i_c \in [1, c]$.

Using the previously defined formulas, we can calculate the interpolated point using

$$\hat{X}(a) = h_1 \times X(a - 1) + h_2 \times X(a + c) + h_3 \times T_1 + h_4 \times T_2. \tag{25}$$

T_1 and T_2 are the tangents at $X(a - 1)$ and $X(a + c)$ and can be estimated using

$$T_1 = \frac{X(a + c) - X(a - 2)}{2} \tag{26}$$
$$T_2 = \frac{X(a + c + 1) - X(a - 1)}{2}. \tag{27}$$

4) Nonlinear Predictive Interpolation: Lippman *et al.* [23] use a predictive interpolation method to replace erroneous data points. In their algorithm, a sequence of m samples before and n samples after the sequence of ectopic beats is defined. This operation results in a sequence of $m + c + n$ samples (where c is the number of consecutive artifacts). We call this sequence the original template (OT) and is obtained as follows:

$$OT = X(a - m), \ldots, X(a + c + n - 1) \tag{28}$$

where a is the index of the first artifact. Now, the middle c points must be replaced by b samples (where b is the number of interpolated points that must be inserted between $X(a - 1)$ and $X(a + c)$). To do that, the m and n points from this sequence are compared with the m and n points from all other sequences of size $m + b + n$. Using the Cartesian distance metric, the closest matching sequence is found. This is called the matched template (MT). The mean of the latter sequence is adjusted so that it matches that of the m and n samples of the OT. Then, the b samples of the MT are inserted into the RR time series between $X(a)$ and $X(a+c)$ in place of the erroneous ones. The study of [23] sets both m and n to 3.

F. Proposed Interpolation Technique

We propose an interpolation technique based on a combination of the cubic and nonlinear predictive interpolation methods. We call the algorithm, cubic predictive interpolation. The algorithm works as follows.

1) Apply the cardinal cubic spline interpolation to calculate the required samples (use the same method provided in Section III-E3).
2) Find the MT, as described in Section III-E4.
3) Calculate $h_1(s)$, $h_2(s)$, $h_3(s)$, and $h_4(s)$ using (20)–(23). In this case, s is defined as

$$s = \frac{i_b}{b + 1} \tag{29}$$

where b is the number of samples to be interpolated and i_b is the index over these samples (i.e., $i_b \in [1, b]$). Therefore, we can calculate the cubic trend (CT) of the MT as follows:

$$CT(i_b) = h_1 \times MT(i_b - 1) + h_2 \times MT(i_b + b) + h_3 \times T_1 + h_4 \times T_2. \tag{30}$$

T_1 and T_2 are the tangents at $MT(i_b - 1)$ and $MT(i_b + c)$ and can be estimated using

$$T_1 = \frac{X(a + c) - X(a - 2)}{2} \tag{31}$$
$$T_2 = \frac{X(a + c + 1) - X(a - 1)}{2}. \tag{32}$$

4) Remove the CT from the MT sequence by subtracting CT from MT. This produces the detrended MT (DMT).
5) Add the b samples of the DMT onto the ones interpolated in step 1.

G. Proposed Artifact Correction Method

All artifact detection techniques listed in Section III-D provide conditions for sample correction. In other words, when a sample is judged to be an artifact, correctional measures are taken. Nonetheless, these measures depend on the nature of the artifact. Table III provides a description of the corrective measures that should be undertaken to correct artifacts belonging to the three categories previously defined.

IV. EVALUATION

A. Artifact Detection Evaluation

1) Evaluation Data Set: We evaluated the above presented algorithms on 79 records containing at least one artifact randomly chosen from the Massachusetts Institute of Technology-Beth Israel Hospital normal sinus rhythm and long-term databases on PhysioNet [30]. Each record is a 5 min segment taken from one of 25 long term ECG recordings. The work presented in this paper focuses on nonpathological artifacts. All records used were manually annotated (or automatically annotated but later visually verified) by experts for ectopic beats, unclassifiable beats, and other potential artifacts. We looked in particular for the following types of annotations:

1) premature atrial contractions;
2) supraventricular premature contractions;
3) premature ventricular contractions;
4) unclassifiable beats;
5) isolated QRS-like artifacts;
6) pauses in measurement.

TABLE III
CORRECTION ACTIONS FOR VARIOUS TYPES OF ARTIFACTS

Category	Condition	Cause	Corrective Action
1	**a)** $X(a) > 2.5 \times \bar{X}$	An interruption in measurement occurred.	The HRV record can no longer be treated as a contiguous record in any analysis exercise. The value I_3 indicates the length of the interruption period.
	b) $X(a) \leq 2.5 \times \bar{X}$	It is likely that $X(a)$ was produced by QRS detection miss.	Interpolate two HRV samples to replace $X(a)$ and $X(a+1)$
2	**a)** $I_3 \approx \bar{X}$	An ECG component (likely a T wave) was mistaken for an R peak.	Remove $X(a)$ and $X(a+1)$ and Insert I_3 in their place
	b) $\bar{X} \ll I_3 \ll 2 \times \bar{X}$	Ectopic beat	Interpolate two HRV samples to replace $X(a)$ and $X(a+1)$
3	All occurrences of this category	Ectopic beat	Interpolate two HRV samples to replace $X(a)$ and $X(a+1)$.

The above presented list compiles the most common types of nonpathological artifacts. The first three types refer to various ectopic beats and the latter three refer to nonectopic artifacts. From the list of annotations accompanying each ECG record, an RR time series is produced that is calculated using the time stamps of the annotated normal beats and the other erroneous features listed above. The following summarizes the categories of the artifacts present in the 79 records used for the test.

1) 12 Category 1 artifacts.
2) 88 Category 2 artifacts.
3) 120 Category 3 artifacts.

2) Finding the Optimal Thresholds: Most of the artifact detection algorithms, we have introduced in this paper include a test based on one or more threshold values to assess the correctness of a sample. Threshold values have been suggested by the works that proposed these algorithms (namely [19], [21], [24]–[26], and [31]). Nonetheless, we opted to find the optimal thresholds to achieve a fair comparison between these algorithms and our proposed one. Consequently, we devised a universal performance indicator against which the effectiveness of threshold values will be determined. The performance indicator considers the rate of accurate detection of artifacts and the rate of false detections. False detections cause the correction of valid RR samples, which negatively affects the signal's integrity. Good threshold values should balance both rates (maximize the true detection and minimize the false detection). Therefore, the performance indicator simply corresponds to the difference between the true positive detection rate (TPDR) and FPDR.

We ran the algorithms on our data in a batch mode, while modifying the threshold values for every single execution. For ST (1), we tested every possible value for $0 \leq R_1 \leq 700$ and $700 \leq R_2 \leq 3000$ (with a step size of 5). We found the optimal values for R_1 and R_2 to be 600 and 900, respectively, yielding a performance indicator value of 50.70%. For MA (3), we tested all values for $0 \leq \theta \leq 1$ and $1 \leq \alpha \leq 3$ (with a step size of 0.05). We found the optimal values for θ and α to be 0.85 and 1.15, respectively, yielding a performance indicator value of 70.28%. Moreover, for RTC (12), we tested all values for $0 \leq m \leq 50$ (with a step size of 1), $10 \leq \min(AT) \leq 250$ and $10 \leq \max(AT) \leq 250$ (with a step size of 10). We found that the best results are obtained when m, $\min(AT)$, and $\max(AT)$ are equal to 7, 60, and 200, respectively (yielding a performance index of 80.76%).

The proposed PWIR algorithm makes use of three thresholds, one for each category of artifacts. Nonetheless, from our preliminary tests, we noticed that the algorithm performs well when $T_1 \approx T_2 \approx T_3 + 1.6$. In addition, similarity, we found a window size of 60 to perform optimally for PWIR and WIR, and therefore ω was fixed to 60 (8) and (9). Using this information, we reduced the amount of thresholds for these algorithms to one. This allows us to compare the PWIR and WIR more easily with the IPFM and IR, both of which rely only on one threshold. We tested all values between 0.1 and 15 (with a step size of 0.1) to find the optimal threshold for algorithms PWIR, IPFM, IR, and WIR. Fig. 8 plots the performance indicator versus the threshold for these algorithms. Therefore, the optimal threshold for each algorithm is associated with the peak data point on the plot
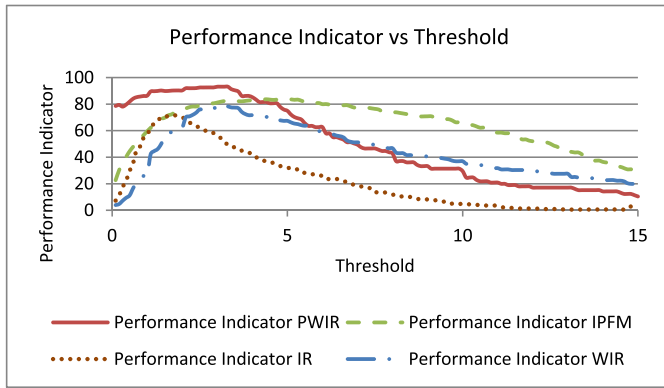
Fig. 8.   Performance indicator plotted against possible threshold values.

TABLE IV
OPTIMAL THRESHOLD VALUES FOR THE IPFM, PWIR, WIR,
AND IR ALGORITHMS

Algorithm	Optimal Threshold(s)
IPFM	5.3
PWIR	3.3
WIR	3.0
IR	1.7
MA	0.85 and 1.15
ST	600 and 900
RTC	7, 60 and 200

for that algorithm. Table IV summarizes the optimal thresholds for all discussed algorithms.

3) Artifact Detection Results and Discussion: Now that we have established the optimal threshold(s) for each algorithm, we can compare the TPDR for each one of them. Table V presents the detection rate of the algorithms with respect to the list of annotations provided in the original ECG records, we obtained from PhysioNet (that is how many of the previously annotated artifacts were detected by the algorithm). The PWIR algorithm produces the highest detection rate; it has detected 93.80% of the artifacts. The WIR and IPFM are somewhat close seconds with rates of 88.1% and 87.62%, respectively. However, in terms of the FPDR, IPFM performs far better than WIR. Nonetheless, the FPDR of the PWIR is still significantly lower than that of the IPFM (0.66% compared with 4.25%). Fig. 9 shows an example of a signal filtered using both IPFM and PWIR. As it can be seen toward the beginning of the record, the IPFM algorithm incorrectly identifies a string of nonerroneous points as artifacts (the dashed square identifies these samples). In addition, the ST exhibits a high FPDR compared to all other algorithms (even though it produced a higher TPDR than Berntson's). Therefore, to fairly rank the algorithms by considering both the TPDR and FPDR, we have to look at their performance index (which we defined as the difference between the TPDR and FPDR). The PWIR algorithm is the most effective at detecting artifacts with a performance index of 93.14%. In addition, interestingly, even though both WIR and IR make use of a median based filter, WIR performs notably better than IR. The improvement in performance can be attributed to the flexibility of the

windowed approach when it comes to dealing with HRV signals resulting from highly dynamic heart rates. The IR algorithm examines the whole record at once, and therefore, in the presence of rapid heart rate change (a common phenomenon during movement in ambulatory settings), the median calculated in (6) is skewed. Conversely, WIR divides the signal into overlapping windows, and thus ensures a more localized calculation of the median. Fig. 10 displays an example of such phenomenon (the dashed circles identify artifacts).

Fig. 11 compares the IPFM, PWIR, WIR, IR, and RTC algorithms more closely by plotting the TPDR against the FPDR over a threshold range of $0.1 \leq T \leq 15$ for IPFM, PWIR, WIR, and IR and a range of $0 \leq m \leq 50$, $10 \leq \min(AT) \leq 250$, and $10 \leq \max(AT) \leq 250$ for RTC. This shows that the TPDR for the PWIR algorithm increases sharply with respect to FPDR compared with the other algorithms that have slower rising curves. Note that the plot of Fig. 11 is purposely cropped at FPDR value of 20, as beyond this false detection rate, an algorithm become barely useful.

Both WIR and IR algorithms are very effective at identifying big impulses in an RR time series that are usually caused by artifacts. Nonetheless, some category 3 artifacts that are caused by ectopic beats tend to display subtle impulses that might be drowned among the other, nonerroneous peaks in the signal (these occur especially in highly variable signals). Fig. 12 shows an example of such scenario where the IR algorithm fails to detect the artifacts, the WIR algorithm detects one pulse and the PWIR catches both pulses.

Note that the latter evaluation was limited to nonpathological HRV artifacts. We did not consider pathological arrhythmias. The works discussed in this paper, including the proposed algorithm, are not effective, or have not been tested, when it comes to handling pathological arrhythmias, such as nocturnal trapezoidally shaped heart rate increases. Sapoznikov *et al.* [33] have presented a study on the subject of pathological arrhythmias and their effect on HRV analysis.

B. Interpolation Evaluation

After detecting the artifacts, the task of minimizing their effect on the integrity of the signal begins. We have proposed the correction procedure presented in Table III. One of the main components of this procedure is the interpolation of new points to replace erroneous ones. We have introduced the cubic predictive interpolation method in Section III-F as our contribution to this cause. To assess its effectiveness in estimating the best possible points that would fit well within the context of the signal with respect to the other existing techniques, we have the devised the following evaluation procedure.

1) Randomly choose 123 artifact free ECG records from the normal sinus rhythm database on PhysioNet (described in Section IV-A1) and compose their corresponding RR time series (each source ECG record is 5 min long).

TABLE V
SUMMARY OF ARTIFACT DETECTION RESULTS

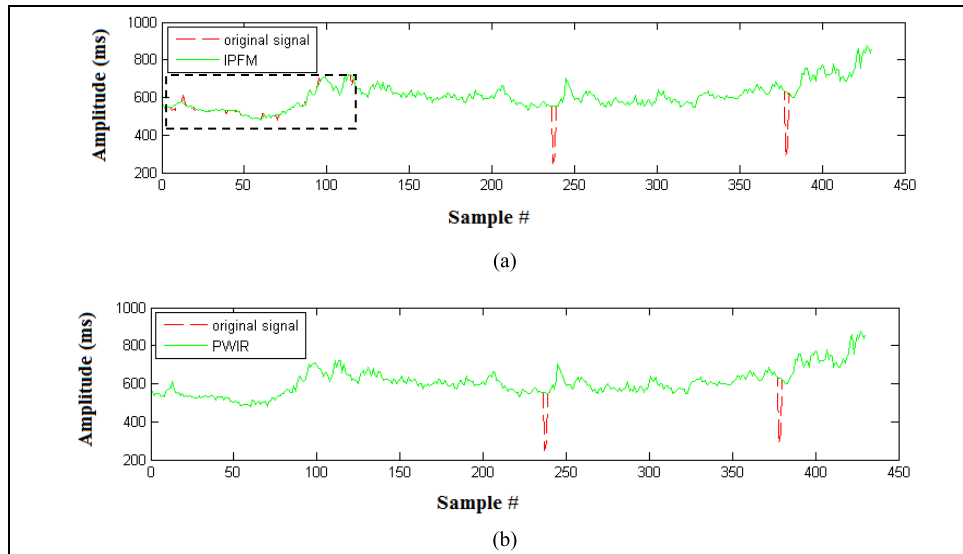|	IPFM	PWIR	WIR	IR	MA	ST	RTC	Berntson's
TPDR	87.62	93.80	88.10	82.04	81.90	80.00	85.71	60.95
FPDR	4.25	0.66	8.56	10.01	11.62	29.30	4.96	0.73
Performance Index	83.37	93.14	79.53	72.03	70.28	50.70	80.76	60.22

Fig. 9. Comparison between the IPFM and PWIR when it comes to FPDR. (a) Integral pulse frequency modulation algorithm. (b) Pattern windowed impulse rejection filter algorithm.
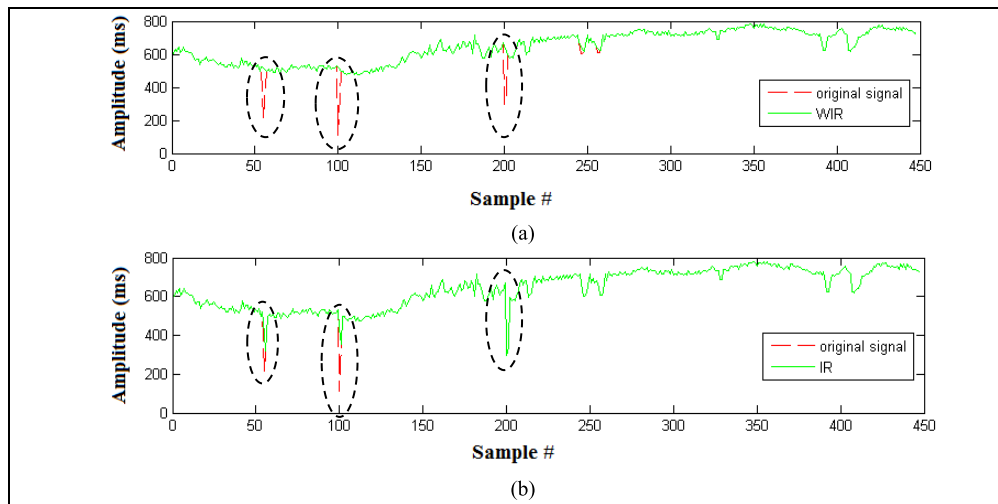

Fig. 10. Comparison between the WIR and IR during quick heart rate changes. (a) Windowed impulse rejection filter algorithm. (b) Impulse rejection filter algorithm.

2) Artificially, introduce artifacts using the method used in [23].
3) Correct the artifacts using the procedure of Table III.
4) Compare the original RR time series to the corrected one by calculating the average error for the following time- and frequency-domain HRV parameters.

   a) LF/HF: low-frequency to high-frequency ratio.
   b) LF: low frequency.
   c) HF: high frequency.
   d) RMSSD: root mean square of successive differences.
   e) Mean R–R: average R to R interval.
   f) SDNN: standard deviation of R to R intervals.

5) In addition to the aforementioned parameters, calculate the root mean square error (RMSE) between the original and corrected signals.

TABLE VI
COMPARISON OF INTERPOLATION TECHNIQUES FOR FREQUENCY-DOMAIN PARAMETERS

		Linear			Median			Nonlinear Predictive			Cubic			Proposed Method		
		1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108
LF/HF	Mean Error %	23.48	7.82	2.35	22.96	8.31	2.52	11.58	6.33	2.43	17.59	6.71	2.13	9.26	5.44	2.19
	SD	0.18	0.08	0.02	0.26	0.10	0.03	0.09	0.07	0.03	0.17	0.07	0.02	0.08	0.05	0.02
LF	Mean Error %	5.83	3.27	1.13	6.93	3.88	1.42	8.48	3.75	1.55	5.54	3.16	1.15	5.74	3.13	1.43
	SD	0.06	0.04	0.01	0.11	0.05	0.02	0.06	0.04	0.02	0.07	0.04	0.01	0.05	0.03	0.01
HF	Mean Error %	19.09	7.08	2.22	17.10	6.80	2.13	11.11	4.96	1.81	13.27	5.38	1.87	7.76	4.03	1.72
	SD	0.08	0.05	0.02	0.09	0.05	0.02	0.07	0.05	0.02	0.07	0.05	0.02	0.07	0.04	0.02

TABLE VII
COMPARISON OF INTERPOLATION TECHNIQUES FOR TIME-DOMAIN PARAMETERS

|		Linear			Median			Nonlinear Predictive			Cubic			Proposed Method		
		1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108	1/12	1/36	1/108
RMSSD	Mean Error %	7.71	2.50	0.70	4.95	1.71	0.56	42.22	9.98	4.45	9.37	2.99	0.84	2.44	1.26	0.54
	SD	0.03	0.02	0.01	0.04	0.02	0.01	0.34	0.17	0.11	0.03	0.02	0.01	0.02	0.02	0.01
Mean R-R	Mean Error %	0.08	0.05	0.03	0.13	0.07	0.04	0.10	0.05	0.03	0.08	0.05	0.03	0.10	0.05	0.03
	SD	0.001	0.001	0.000	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001	0.000
SDNN	Mean Error %	2.16	0.83	0.27	1.77	0.78	0.28	5.48	1.47	0.80	1.74	0.74	0.27	1.48	0.66	0.32
	SD	0.02	0.01	0.00	0.02	0.01	0.00	0.08	0.02	0.02	0.02	0.01	0.00	0.01	0.01	0.00
RMSE	Mean Error %	14.38	8.13	4.15	16.51	9.46	4.90	22.96	11.04	5.99	13.52	7.69	3.94	15.57	8.36	4.29
	SD	14.92	8.67	4.56	16.69	9.84	5.32	10.96	7.49	5.28	13.47	8.00	4.33	12.28	7.23	3.73
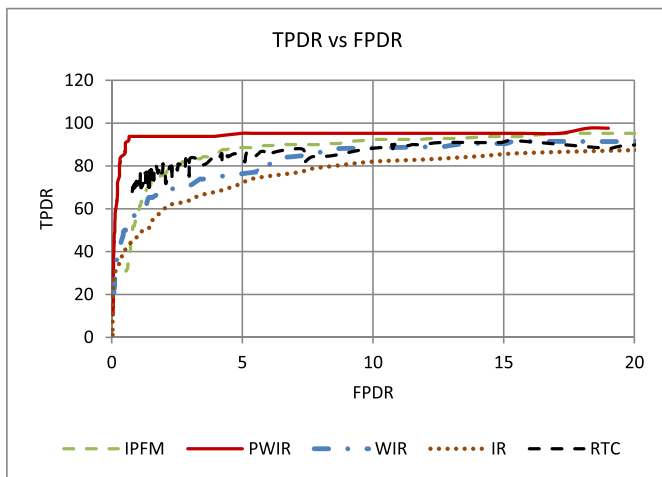
Fig. 11. TPDR versus FPDR for the IPFM, PWIR, WIR, IR, and RTC algorithms.

1) Interpolation Results and Discussion: We ran the correction procedure using the various interpolation algorithms for all the RR time series. We produced an average RMSE and average percentage of error for each HRV parameter discussed above for various percentage of artificially introduced artifacts [(1/12)%, (1/36)%, and (1/108)%]. The frequency-domain parameters were calculated using Welch's method (window size of 256, 50% overlap and 4 Hz sampling rate). Table VI summarizes our frequency-domain results, while Table VII presents our time domain results. For the smallest percentage of artifacts, the average error for the frequency-domain parameters for all algorithms is fairly comparable, with the cubic, nonlinear predictive, and cubic predictive interpolation methods slightly edging the other ones. The cubic interpolation produces an exceptionally low average error for LF. Reducing the error on the low frequencies is a definite strength of the cubic interpolation which produces smooth curves mimicking very well the low-frequency sinusoids of the original signal. Nonetheless, using such approach, the high-frequency sinusoids of the original signal are not considered. In terms of the time-domain parameters, the nonlinear predictive approach produces a very high average error for RMSSD. The calculation of RMSSD involves adding the difference between successive samples. When the nonlinear predictive method is used, the difference of successive samples can differ significantly between the original and interpolated signal, especially at the boundaries of the interpolated sequence (i.e., difference between the sample just before the point of insertion and the first interpolated sample and between the last interpolated sample and the one that follows). All other algorithms incorporate in a more
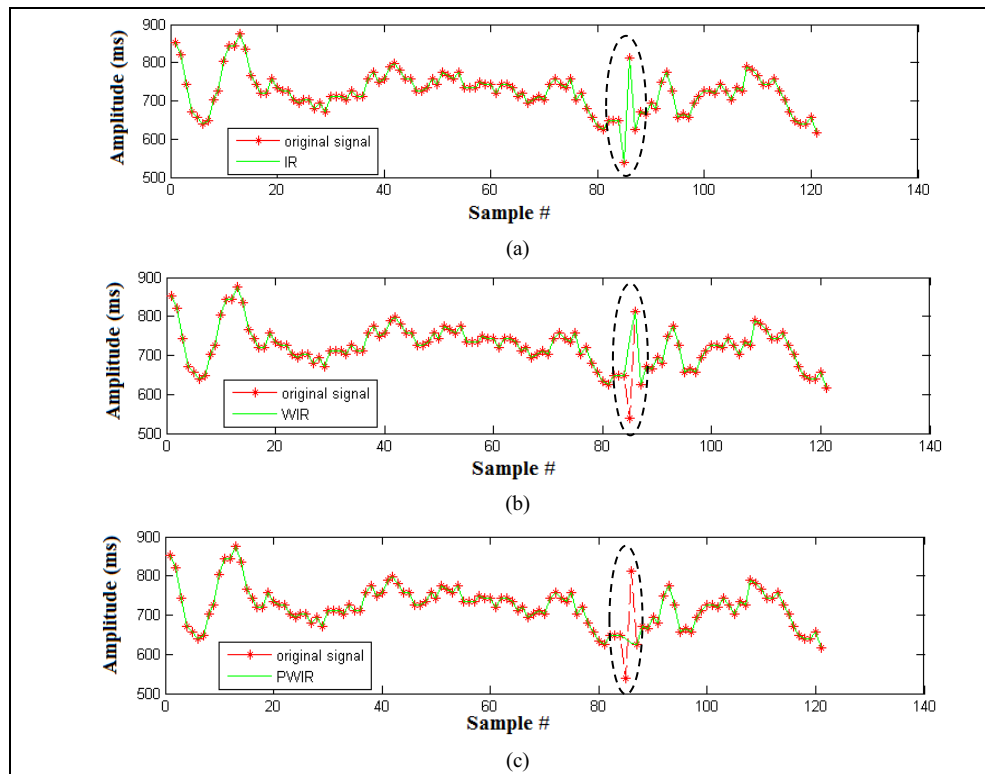
Fig. 12.   Subtle category 3 impulses caused by a premature ventricular contraction. (a) Impulse rejection filter algorithm. (b) Windowed impulse rejection filter algorithm. (c) Pattern windowed impulse rejection filter algorithm.

effective way the samples surrounding the interpolated area to produce a less erroneous RMSSD. As the percentage of artifacts increases, we start to clearly observe the weakness of the cubic interpolation in terms of producing a large average error on the HF parameter. In addition, we start to observe a clear advantage for the proposed method over all of the other algorithms on the average error for all of the frequency- and time-domain parameters being evaluated. Even at the very high artifacts percentage of $(1/12)\%$, the proposed method produces an average error below 3% for all HRV time-domain parameters and below 10% for all frequency-domain HRV parameters.

V. CONCLUSION

In this paper, we have presented the PWIR algorithm for artifact detection. We have compared our method to seven existing ones (ST, MA, impulse filter, WIR Filter, IPFM, RTC, and Berntson's) and demonstrated its superiority in terms of achieving a high true detection rate (93.80%) and a low false detection rate (0.66%) on our data set. Moreover, we have proposed a procedure for correcting artifacts based on the deletion, insertion, and interpolation of samples whenever needed. We paid particular attention to the interpolation operation and presented a technique that produces more optimal sample estimation than the linear, cubic, nonlinear predictive, and median interpolation methods proposed for the same purpose. Our future work includes the application of the PWIR artifact detection techniques to biofeedback systems responsible for relaying the level of mental stress assessed from an HRV signal to the user as part of a stress management process.

REFERENCES

[1] T. Raita-Aho, T. Saramaki, and O. Vainio, "A digital filter chip for ECG signal processing," *IEEE Trans. Instrum. Meas.*, vol. 43, no. 4, pp. 644–649, Aug. 1994.

[2] S. Ahmad, M. Bolic, H. Dajani, V. Groza, L. Batkin, and S. Rajan, "Measurement of heart rate variability using an oscillometric blood pressure monitor," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2575–2590, Oct. 2010.

[3] M. Chen, O. Boric-Lubecke, and V. M. Lubecke, "0.5-μm CMOS implementation of analog heart-rate extraction with a robust peak detector," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 4, pp. 690–698, Apr. 2008.

[4] E. H. Hon and S. T. Lee, "Electronic evaluations of the fetal heart rate patterns preceding fetal death: Further observations," *Amer. J. Obstetrics Gynecol.*, vol. 87, pp. 814–826, 1965.

[5] H. M. Colhoun, D. P. Francis, M. B. Rubens, S. R. Underwood, and J. H. Fuller, "The association of heart-rate variability with cardiovascular risk factors and coronary artery calcification: A study in type 1 diabetic patients and the general population," *Diabetes Care*, vol. 24, no. 6, pp. 1108–1114, Jun. 2001.

[6] M. Malik *et al.*, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Psychophysiology*, vol. 39, no. 4, pp. 427–436, 2009.

[7] A. Haensel, P. J. Mills, R. A. Nelesen, M. G. Ziegler, and J. E. Dimsdale, "The relationship between heart rate variability and inflammatory markers in cardiovascular diseases," *Psychoneuroendocrinology*, vol. 33, no. 10, pp. 1305–1312, 2008.

[8] S. Boonnithi and S. Phongsuphap, "Comparison of heart rate variability measures for mental stress detection," in *Proc. Comput. Cardiol.*, Sep. 2011, pp. 85–88.

[9] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task force of the European society of cardiology and the North American society of pacing and electrophysiology," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[10] C. De Capua, A. Meduri, and R. Morello, "A smart ECG measurement system based on web-service-oriented architecture for telemedicine applications," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2530–2538, Oct. 2010.

[11] Zephyr. *BioHarness BT.* [Online]. Available: http://www.zephyranywhere.com/products/bioharness-3, accessed Aug. 19, 2013.

[12] Biopeak. *BioFusion PSM.* [Online]. Available: http://biopeak.com/biofusion/, accessed Aug. 19, 2013.

[13] S. Mitra, M. Mitra, and B. B. Chaudhuri, "A rough-set-based inference engine for ECG classification," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 6, pp. 2198–2206, Dec. 2006.

[14] S. Graja and J.-M. Boucher, "Hidden Markov tree model applied to ECG delineation," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2163–2168, Dec. 2005.

[15] T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 4, pp. 1224–1231, Aug. 2003.

[16] N. Hjortskov, D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work," *Eur. J. Appl. Physiol.*, vol. 92, nos. 1–2, pp. 84–89, 2004.

[17] R. Colombo, G. Mazzuero, F. Soffiantino, M. Ardizzoia, and G. Minuco, "A comprehensive PC solution to heart rate variability analysis in mental stress," in *Proc. Comput. Cardiol.*, Sep. 1989, pp. 475–478.

[18] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: A review," *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, 2006.

[19] J. McNames, T. Thong, and M. Aboy, "Impulse rejection filter for artifact removal in spectral analysis of biomedical signals," in *Proc. 26th Annu. Int. Conf. IEEE EMBS*, Sep. 2004, pp. 145–148.

[20] J. E. Mietus, "Time domain measures: From variance to pNNx," Ph.D. dissertation, Beth Israel Deaconess Med. Center, Harvard Med. School, Boston, MA, USA, 2006.

[21] L. J. M. Mulder, "Measurement and analysis methods of heart rate and respiration for use in applied environments," *Biol. Psychol.*, vol. 34, nos. 2–3, pp. 205–236, Nov. 1992.

[22] M.-Y. Lee and S.-N. Yu, "Improving discriminality in heart rate variability analysis using simple artifact and trend removal preprocessors," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug./Sep. 2010, pp. 4574–4577.

[23] N. Lippman, K. M. Stein, and B. B. Lerman, "Comparison of methods for removal of ectopy in measurement of heart rate variability," *Amer. J. Physiol.*, vol. 267, no. 1, pp. H411–H418, 1994.

[24] J. Mateo and P. Laguna, "Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 334–343, Mar. 2003.

[25] H. Al Osman, M. Eid, and A. El Saddik, "A windowed impulse rejection filter for HRV artifact detection," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2013, pp. 6–11.

[26] J. H. Houtveen and P. C. M. Molenaar, "Comparison between the Fourier and wavelet methods of spectral analysis applied to stationary and nonstationary heart period data," *Physiological*, vol. 38, no. 5, pp. 729–735, 2001.

[27] F. Chen and Y. T. Zhang, "An efficient algorithm to reconstruct heart rate signal based on an IPFM model for the spectral analysis of HRV," in *Proc. 27th Annu. Int. Conf. Eng. Med. Biol. Soc. (IEEE-EMBS)*, Jan. 2006, pp. 936–939.

[28] S. R. Seydnejad and R. I. Kitney, "Real-time heart rate variability extraction using the Kaiser window," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 10, pp. 990–1005, Oct. 1997.

[29] M. A. García-González, C. Vázquez-Seisdedos, and R. Pallàs-Areny, "Variations in breathing patterns increase low frequency contents in HRV spectra," *J. Physiol. Meas.*, vol. 21, no. 3, pp. 417–423, 2000.

[30] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000. [Online]. Available: http://circ.ahajournals.org/cgi/content/full/101/23/e215

[31] J. Rand, A. Hoover, S. Fishel, J. Moss, J. Pappas, and E. Muth, "Real-time correction of heart interbeat intervals," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 946–950, May 2007.

[32] G. G. Berntson, K. S. Quigley, J. F. Jang, and S. T. Boysen, "An approach to artifact identification: Application to heart period data," *Psychophysiology*, vol. 27, no. 5, pp. 586–598, Sep. 1990.

[33] D. Sapoznikov, M. H. Luria, Y. Mahler, and M. S. Gotsman, "Computer processing of artifact and arrhythmias in heart rate variability analysis," *Comput. Methods Programs Biomed.*, vol. 39, nos. 1–2, pp. 75–84, Sep. 1992.

Hussein Al Osman received the B.A.Sc. (Hons.) (*summa cum laude*) degree in computer engineering, the M.A.Sc. degree in electrical engineering, and the Ph.D. degree in electrical engineering from the University of Ottawa, Ottawa, ON, Canada, in 2007, 2009, and 2014, respectively.

He is currently an Assistant Professor of Software Engineering with the School of Electrical Engineering and Computer Science, University of Ottawa. His current research interests include health informatics, serious games for health, affective computing, human–computer interaction, and active biometrics.

Dr. Al Osman has received numerous scholarships and awards, such as the 3X NSERC Scholarships, the Queen Elizabeth II Graduate Scholarship, the Best Paper Award at the DS-RT 2008 Conference, and the Part Time Professor Award.

Mohamad Eid received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Ottawa, ON, Canada, in 2010.

He was a Teaching and Research Associate with the University of Ottawa from 2008 to 2012. He is currently an Assistant Professor of Practice of Electrical Engineering with the Division of Engineering, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates. He has co-authored a book entitled *Haptics Technologies: Bringing Touch to Multimedia* (Springer, 2011). His current research interests include multimedia haptics, tangible human–computer interaction, biofeedback, and instrumentations (sensors and actuators).

Dr. Eid was a recipient of the Best Paper Award of the DS-RT Conference in 2008 and the prestigious ACM Multimedia Grand Challenge Most Entertaining Award for *HugMe: Synchronous Haptic Teleconferencing System* in 2009. He is the Co-Chair of the 3rd International IEEE Workshop on Multimedia Services and Technologies for E-health in 2013, the Technical Chair of the Haptic-Audio-Visual Environment and Gaming Workshop in 2013, and a Steering Committee Member for over 15 conferences.

Abdulmotaleb El Saddik (M'01–SM'04–F'09) is currently a Distinguished University Professor and the University Research Chair of the School of Electrical Engineering and Computer Science with the University of Ottawa, Ottawa, ON, Canada. He is a Global Visiting Professor of Engineering with the Division of Engineering, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, and an Honorary Professor with Sichuan University, Chengdu, China. He is an Internationally Recognized Scholar who has made strong contributions to the knowledge and understanding of multimedia computing, communications, and applications. He has authored or co-authored four books and over 450 publications, and supervised over 100 researchers.

Prof. El Saddik is also a fellow of the Engineering Institute of Canada and the Canadian Academy of Engineers. He received several international awards, such as the ACM Distinguished Scientist Award and the IEEE Canada Computer Medal. He has chaired over 40 conferences and workshops, and received research grants and contracts totalling over $18 Mio.