IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY SECTION

Received January 28, 2021, accepted March 16, 2021, date of publication April 2, 2021, date of current version April 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3070604

R-R Interval Estimation for Wearable Electrocardiogram Based on Single Complex Wavelet Filtering and Morphology-Based Peak Selection

SUEHIRO SHIMAUCHI[1], (Senior Member, IEEE), KANA EGUCHI[2], (Member, IEEE), RYOSUKE AOKI[2], MASAHIRO FUKUI[1], (Member, IEEE), AND NOBORU HARADA[1], (Senior Member, IEEE)

[1]NTT Media Intelligence Laboratories, Nippon Telegraph and Telephone Corporation, Tokyo 180-8585, Japan
[2]NTT Service Evolution Laboratories, Nippon Telegraph and Telephone Corporation, Kanagawa 239-0847, Japan

Corresponding author: Kana Eguchi (kana.eguchi.gh@hco.ntt.co.jp)

ABSTRACT Recent innovations in wearable electrocardiogram (ECG) devices have enabled various personal healthcare applications based on heart rate variability (HRV). However, wearable ECGs rarely undergo visual inspection by medical experts, hence may contain noise and artifacts. Because apparent changes in the recorded ECGs caused by noise and artifacts may hamper the extraction of QRS complexes, an R-R interval (RRI) estimation algorithm tolerant to these measurement faults is required as the initial step toward HRV analysis using wearable ECGs. This paper proposes a semi-real-time RRI estimation for wearable ECGs utilizing a two-stage structure. In the preprocessing stage, we use a complex-valued wavelet that can adaptively fit to morphological variations of the QRS complex while retaining computing resources for extracting the QRS complex features. In the decision stage, we make use of complex-valued features and select appropriate QRS complexes in consideration of three features: peak magnitude, peak location, and peak morphology (phase). Initial evaluations show that the QRS complex detection performance of the proposed method achieved the F1 score of 0.952 ± 0.040 when targeting pseudo ECG data created from open data assuming wearable ECGs, and of 0.986 ± 0.018 when targeting actual ECG data recorded by a shirt-type wearable ECG device during an exercise activity. Furthermore, the proposed method was able to suppress overlook or misdetection of QRS complexes, so the obtained RRIs are closer to the reference RRIs. The proposed method therefore contributes to achieving accurate HRV analysis using wearable ECGs in terms of obtaining accurate RRIs.

INDEX TERMS Electrocardiogram (ECG), heart rate variability (HRV), QRS complex detection, R-R interval (RRI), wavelet analysis, wearable device.

I. INTRODUCTION

Recent innovations in wearable electrocardiogram (ECG) devices have enabled the provision of various personal healthcare applications based on heart rate variability (HRV) [1] calculated from a continuous ECG record, such as sleep monitoring [2] and long-term monitoring for rehabilitation [3] or driver drowsiness [4]. With long-term ECG recording for HRV analysis, the device should be easy-to-use for non-experts and have a minimal effect on the user's daily life. In this sense, a shirt-type wearable ECG device with embedded measurement electrodes and lead wires has attracted attention recently [5]–[7]. With this device, users simply wear a special shirt with embedded measurement electrodes and lead wires and attach a small dedicated device for the ECG recording.

Compared to ECGs recorded by in-hospital ECG devices, however, those recorded by wearable devices are generally more susceptible to noise and artifacts due to external factors such as body movements, respiration, and perspiration [8], [9]. Because wearable ECGs recorded for non-clinical healthcare applications are visually inspected by

The associate editor coordinating the review of this manuscript and approving it for publication was Emil Jovanov.

well-trained medical experts only rarely, noise and artifacts may remain and hamper the extraction of important ECG features by causing apparent changes in the recorded ECGs even if the positions of the measurement electrodes have not significantly changed. Furthermore, when using the aforementioned shirt-type wearable ECG device, the recorded ECGs are also affected by the characteristics inherent in clothing: for example, a shirt deforms itself along with body movement, and this may cause deformation of the embedded measurement electrodes, inducing potential displacements between the measurement electrodes and the skin surface that ultimately results in impedance fluctuations [8], which we can observe as noise or artifacts. In addition, the position of the measurement electrodes themselves is affected by both the way of wearing and the person's physique, even in the same subject, and the change of the measurement electrode position can cause morphological changes in the recorded ECGs in theory.

Assuming a situation in which non-experts are asked to record ECGs during daily life, ideally, the ECG features should be extracted while suppressing the effect of noise and artifacts, as we would prefer not to have non-experts manually reposition the location of measurement electrodes all the time in daily life. In other words, robustness against morphology variations and artifacts is required. Furthermore, the ECG feature extraction should be accomplished with a limited dataset and limited computing resources (i.e., a mobile environment such as a smartphone) in real-time or semi-real-time (i.e., with only slight delay), because the currently available shirt-type ECG devices generally transfer the recorded ECGs by means of real-time/semi-real-time processing using the combination of a dedicated wireless transmitter and a commercial smartphone/tablet [10], both of which have limited storage and computing resources compared to a personal computer.

The ECG signal waveform is commonly modeled as shown in Fig. 1. The R-R interval (RRI), which is defined as the interval of two adjacent local maxima of QRS complexes (i.e., R waves), is one of the most important ECG features, especially from the perspective of HRV analysis. QRS complex detection has therefore been a major research topic in the ECG signal processing field for several decades [11]–[13]. The general scheme for QRS complex detection is built on a two-stage structure [12], [14]: a preprocessing stage comprising linear filtering and nonlinear filtering for feature extraction or denoising, and a decision stage comprising peak detection logic and the decision for selecting applicable QRS complexes alone.

The typical preprocessing scheme is based on a multi-scale signal decomposition such as wavelet transform (WT) or empirical mode decomposition (EMD). The difference between the two is that WT uses a set of pre-fixed filters whereas EMD adaptively changes the filters in a data-driven way [15]. WT has many variations, which can be categorized by the difference of their transformation structure or the wavelet functions [16]. Discrete wavelet transform (DWT)

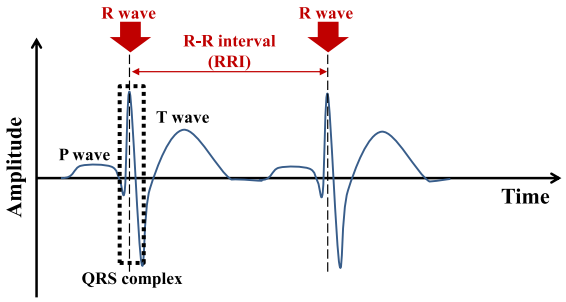

FIGURE 1. ECG waveform and R-R interval (RRI).

is used mainly for data compression purposes [17], [18] and compactly represents the information in the maximally decimated scale-and-shift domain with the orthogonal wavelet bases. Continuous wavelet transform (CWT) is also a useful candidate for denoising or feature-extraction purposes [19], which can be, in practice, implemented with discretized approximation. Because CWT can use even non-orthogonal wavelet bases, the choice of the mother wavelet function for CWT is more flexible than that for DWT. For ECG feature extraction, typical wavelets include the Mexican hat and the Morlet wavelet [20], which are expected to have a similar waveform morphology to the ECG signal. The EMD approach, on the other hand, is more appropriate for cases of extracting a signal whose waveform can vary widely in time-scale: for example, EMD has outperformed WT on an experimental task of extracting the respiratory signal from observed ECG signals [15]. In general, however, all these multi-scale signal decomposition approaches would require high computational cost in exchange for traceability of variations in the target signal features.

Regarding the decision stage, each QRS complex location is estimated from the feature sequence extracted in the preprocessing stage. In the feature domain of interest, the QRS complex parts are assumed to dominate other parts of the ECG signal. Most of the localization methods are therefore based on a fixed or variable thresholding in terms of magnitude [11], [21]. However, this magnitude thresholding is not necessarily effective when it comes to selecting applicable QRS complexes: fixed thresholding may cause overlook or overdetection depending on the setting value, whereas variable thresholding may sometimes fail to follow the change and result in overlook due to the changed threshold [22]. Although the recently proposed time-attention method [23] predicts the probabilities of the regions of QRS complexes by using a convolutional neural network (CNN), this approach may not be free from the aforementioned thresholding issue because its final decision is based on a variable threshold, and it also requires a large amount of computation.

In this paper, we propose a semi-real-time RRI estimation for wearable ECG based on a simple and robust two-stage structure. For the preprocessing stage, we use a single wavelet bandpass filter for the QRS-complex feature extraction instead of multiple filter banks such as the WT or EMD. Similar approaches to the computational simplicity can be

found in previous studies: Pan and Tompkins [11] used a recursive bandpass filter with integer coefficients, whereas Gul *et al.* [24] proposed selecting only a single-scale output of tree-structured wavelet filterbanks as a feature of interest. Although the filters used in [11], [24] are real-valued, we herein use a complex-valued wavelet as a filter, as it can adaptively fit to morphological variations of the QRS complex while retaining computing resources, especially for QRS complexes undergoing morphological changes due to noise and artifacts. When recording ECGs in a daily life environment, we often encounter artifacts with a larger magnitude than those of the QRS complexes, which means that QRS complex detection based only on the threshold is not always effective in terms of improving robustness. For the decision stage, we therefore make use of complex-valued features and select appropriate QRS complexes in consideration of three features, namely, peak magnitude, peak location, and peak morphology (phase). Note that our proposed method is essentially based on our previous study [25], in which multi-scale complex-valued wavelets are utilized, but we focus here on appropriate RRI estimation for HRV analysis under the actual environment; therefore, we only utilize a single complex-valued wavelet as a preprocessing filter to reduce computational cost and achieve semi-real-time processing. In addition, we evaluate the performance of the proposed method targeting wearable ECGs by using pseudo ECG data created from open data in which we assume the ECGs recorded by shirt-type wearable ECG devices and actual ECG data recorded by a shirt-type wearable ECG device during an exercise activity.

II. RRI ESTIMATION METHOD

Our proposed method is based on the two-stage structure scheme described below. The first stage extracts the QRS complex features from the ECG signal by using a single bandpass filtering with a complex-valued wavelet. The second stage detects the peaks of the complex valued features and selects appropriate peaks among them based on a criterion that takes into account the location, magnitude, and morphology (phase) of the detected peaks.

A. SINGLE COMPLEX-VALUED WAVELET FOR FEATURE EXTRACTION

In the feature extraction stage, a bandpass filtering is applied by choosing a specific single-scale output of the CWT [16]. A general form of the CWT can be described as

$$X(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\, \psi^* \left(\frac{t-b}{a} \right) dt \tag{1}$$

where $x(t)$ is a continuous-time ECG signal, $\psi(t)$ is a wavelet function, $X(a,b)$ is the wavelet transform at scale a and shift b of the wavelet, and * indicates the complex conjugate.

As a mother wavelet, we choose the following complex-valued function:

$$\psi_{\mathrm{HWCS}}(t) = \begin{cases} \frac{1}{2}\left(1+\cos\frac{\omega_0}{2}t\right) e^{j\omega_0 t}, & \text{if } |t| < \frac{2\pi}{\omega_0} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

This is a Hann-windowed complex sinusoid (HWCS), which is often used in short-time Fourier transform (STFT) and was also used in a previous study [26]. In WT applications, although the Morlet wavelet is often used as a complex-valued wavelet [27], we consider that the HWCS wavelet is more suitable for extracting the feature with low oscillations that can be observed in the QRS complex waveform morphology. There are two versions of the Morlet function:

$$\psi_{\mathrm{Morlet}}(t) = \frac{1}{\sqrt[4]{\pi}} e^{-\frac{t^2}{2}} e^{j\omega_0 t} \tag{3}$$

and

$$\psi_{\mathrm{C\_Morlet}}(t) = \frac{1}{\sqrt[4]{\pi}} e^{-\frac{t^2}{2}} \left(e^{j\omega_0 t} - e^{-\frac{\omega_0^2}{2}} \right) \tag{4}$$

where the function in (4) is referred to as a complete Morlet wavelet. The properties of (2), (3), and (4) are shown in Fig. 2, where $\omega_0 = 1.9635$ is given for all functions for comparison in a low-oscillation case. Fig. 2(d) shows that the frequency characteristics of the HWCS and the complete Morlet satisfy the admissibility condition [16], i.e., their spectral magnitudes at zero frequency are close to zero. Moreover, the HWCS can be a better analytic wavelet than the complete Morlet because its spectral leakage at the negative frequencies is smaller.

The main reason we chose the complex-valued wavelet here is its flexibility to changes in waveform morphology. Fig. 3 shows an example of the variations in QRS complex patterns of chest leads V1 to V6, which are attached at different chest positions during a multi-lead ECG measurement by in-hospital ECG devices [28]. Although these variations are originally observed by placing the measurement electrode on the designated position corresponding to the desired chest read, similar apparent changes in QRS complex can often be observed in wearable situations: even though the positions of the measurement electrodes are comparatively stable, contaminated noise or artifacts cause apparent changes in the shape of QRS complexes. Furthermore, for wearable ECG with an electrode embedded inside a shirt, it is difficult to attach the electrode at the same position every time or for every person who wears it. The complex wavelet can flexibly adapt to such waveform variations by changing its phase parameter. Fig. 4 shows the variations in waveforms synthesized from the HWCS wavelet with different phases, $\phi = 0$ to π.

For practical implementation, we choose a specific scale $a = 1$ for the HWCS wavelet with $\omega_0/2\pi = 20$ Hz, in accordance with the typical QRS complex frequency range [29]. Then, a discrete version of the HWCS wavelet is implemented

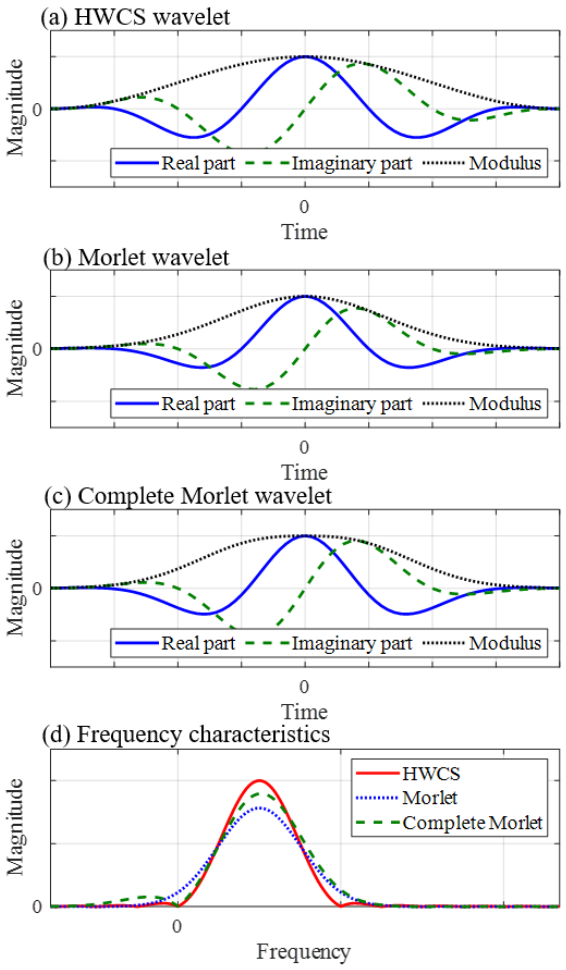

FIGURE 2. Comparison of complex wavelets.

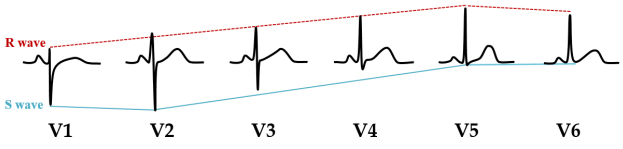

FIGURE 3. Variations in QRS complex patterns of chest leads from V1 to V6 (This set of ECG waveforms originally appeared in [28], on which we drew additional lines indicating R wave and S wave).

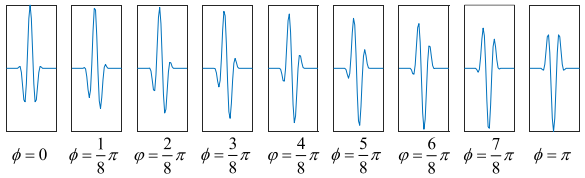

FIGURE 4. Variations in waveforms synthesized from HWCS wavelet with different phases 0 to π.

as a transversal digital filter for the discrete-time ECG signal $x(n)$:

$$X(n) = \mathbf{\Psi}^H \mathbf{x}(n) \tag{5}$$

where

$$\mathbf{x}(n) = [x(n-N), \cdots, x(n+N)]^T \tag{6}$$

$$\mathbf{\Psi} = \left[\psi^*_{\mathrm{HWCS}}\left(\frac{-N}{f_s}\right), \cdots, \psi^*_{\mathrm{HWCS}}\left(\frac{N}{f_s}\right)\right]^H \tag{7}$$

^{T} and ^{H} indicate the transposition and the conjugate transposition, respectively, $2N+1$ is the vector dimension and corresponds to the truncation range of the wavelet function, and n is a discrete time index sampled at the sampling frequency f_s. The filter output $X(n)$ is calculated at every n without decimation.

B. PEAK DETECTION AND SELECTION

The filter output $X(n)$ corresponds to the complex-valued feature sequence of the QRS complexes. By searching for n_p, i.e., the location of the p-th peak of the modulus of the feature sequence, we can obtain the candidates of the location of each QRS complex. Some of the peaks are due to artifacts, and their magnitudes might be larger than those of the QRS complexes in wearable EGC measurements. Therefore, magnitude-based thresholding is not applicable in this case. As the complex-valued feature coefficients provide the morphology information from their phase properties, the difference in the morphologies around the peaks can also be taken into account.

In order to judge whether the detected peak location n_p is that of a QRS complex peak or not, the following semi-real-time peak selection criteria is defined:

$$C(n_p) = D_{\mathrm{int}}(n_p) + D_{\mathrm{mod}}(n_p) + D_{\mathrm{phs}}(n_p) \tag{8}$$

where

$$D_{\mathrm{int}}(n_p) \propto \left|\left[n_p - \hat{n}_{\mathrm{QRS}(m-1)}\right] - \overline{\mathrm{RRI}}_{m-1}\right|^2 \tag{9}$$

$$D_{\mathrm{mod}}(n_p) \propto \left|\log|X(n_p)| - \overline{\log \mathrm{M}}_{m-1}\right|^2 \tag{10}$$

$$D_{\mathrm{phs}}(n_p) \propto \left|\log\left(\frac{X(n_p)}{|X(n_p)|}e^{-j\bar{\phi}_{m-1}}\right)\right|^2 \tag{11}$$

$D_{\mathrm{int}}(n_p)$ corresponds to a distance measure of the peak intervals between the peak at n_p and the last estimated location of the QRS complex, $\hat{n}_{\mathrm{QRS}(m-1)}$, compared to the short-time-averaged RRI estimate, $\overline{\mathrm{RRI}}_{m-1}$. $D_{\mathrm{mod}}(n_p)$ corresponds to a distance measure of the log modulus of $X(n_p)$ compared to the short-time-averaged log modulus of the estimated QRS complex peaks, $\overline{\log \mathrm{M}}_{m-1}$. $D_{\mathrm{phs}}(n_p)$ corresponds to a distance measure of the phase of $X(n_p)$ compared to the short-time-averaged phase of the estimated QRS complex peaks, $\bar{\phi}_{m-1}$. The m-th estimated location of the QRS complex peak, $\hat{n}_{\mathrm{QRS}(m)}$, can be given by

$$\hat{n}_{\mathrm{QRS}(m)} = \underset{n_p \in S(m)}{\mathrm{argmin}}\, C(n_p) \tag{12}$$

where $S(m)$ is the m-th subset of the feature-sequence peaks detected after $\hat{n}_{\mathrm{QRS}}(m-1)$. For the semi-real-time processing, the range of $S(m)$ should be limited to a few seconds right

after $\hat{n}_{\mathrm{QRS}}(m-1)$. When $\hat{n}_{\mathrm{QRS}}(m)$ is selected, the short-time-averaged parameters are updated as follows.

$$\overline{\mathrm{RRI}}_m = \alpha\overline{\mathrm{RRI}}_{m-1}+(1-\alpha)\left[\hat{n}_{\mathrm{QRS}(m)} - \hat{n}_{\mathrm{QRS}(m-1)}\right] \quad (13)$$

$$\overline{\mathrm{logM}}_m = \beta\overline{\mathrm{logM}}_{m-1} + (1-\beta)\log\left|X\left(\hat{n}_{\mathrm{QRS}(m)}\right)\right| \quad (14)$$

$$\overline{\phi}_m = \gamma\overline{\phi}_{m-1} + (1-\gamma)\angle X\left(\hat{n}_{\mathrm{QRS}(m)}\right) \quad (15)$$

where α, β, γ are time-averaging constants whose values are chosen between 0 and 1.

III. EVALUATION

To evaluate the expected performance in QRS complex detection when targeting ECGs recorded in the daily life environment, we conducted two evaluations targeting different ECG data: experiment 1, targeting pseudo ECG data created using open data [30], in which we assumed the ECGs recorded by shirt-type wearable ECG devices, and experiment 2, targeting actual ECG data recorded by a shirt-type wearable ECG device during an exercise activity.

Because we assume that the expected performance in QRS complex detection can be validated in terms of the appropriateness of detected QRS complexes and the tolerance of QRS complex detection performance to measurement faults (e.g., during/after measurement faults), we compared the performance of the proposed method against conventional methods from the following two perspectives: performance of the QRS complex detection itself, and performance of the RRI calculation as the initial step of HRV analysis. To evaluate the former perspective, we calculated the F1 score [31], [32], which can consider both misdetection and overlook, in addition to the precision and recall utilized in the previous study [33]. All measurements were calculated on the basis of the detected QRS complexes against the reference QRS complexes. The method for calculating the F1 score in each experiment is described in the corresponding subsections. For the latter perspective, we used the tachogram of RRIs and compared the RRIs calculated from the detected QRS complexes against those obtained from the reference RRIs.

To evaluate the performance difference in the QRS complex detection and the RRI calculation derived from the peak detection logic in the ''decision stage'' utilized in the QRS complex detection algorithm, we compared the following three methods in each experiment: (a) the proposed combination, which consists of the proposed wavelet and proposed peak detection, (b) a method that consists of the proposed wavelet and the conventional peak detection logic, and (c) a benchmark method. Because methods (a) and (b) utilize the same wavelet while using different peak detection logic, we expect that the comparative evaluation between these two will clarify the performance difference derived from the decision stage. As the conventional peak detection in method (b), we used the ''findpeaks'' function in MATLAB (The MathWorks, Inc., Natick, MA, USA) and set the ''MinPeakDistance'' value to 250 ms. As the benchmark method, we used the Pan-Tompkins algorithm (PTA) [11], which is also utilized for a gold-standard HRV analysis software [34].

All experiments were conducted after preparing target ECG data in the storage medium of a personal computer (CPU, Intel ® Core™ i7-7700 3.60GHz; RAM, 32.0 GB; OS, Windows 10) that was analyzed off-line afterwards.

A. EXPERIMENT 1: EVALUATION OF QRS COMPLEX DETECTION PERFORMANCE ON PSEUDO ECG DATA CREATED BY OPEN DATA ASSUMING DAILY LIFE ACTIVITIES

In actual ECGs recorded by shirt-type wearable ECG devices during daily life activities, noise and artifacts may randomly occur at any time due to measurement faults [9]. Furthermore, identical noise or artifacts will not be repeatedly observed even when the same user performs the same movements because these measurement faults are also affected by the electrode conditions (e.g., wet/dry, placement position, contact area, flexion of electrode itself, or friction between electrode and skin surface), which can hardly be exactly the same every time.

We therefore first evaluated the performance of the proposed method targeting pseudo ECGs that contain noise or artifacts similar to those observed in the ECGs recorded by shirt-type wearable ECG devices in a daily life environment. As the evaluation reference, we used the annotations of QRS complexes that came with the corresponding ECGs in the open dataset.

1) TARGET PSEUDO ECG DATA

The main factors affecting the accuracy of QRS complex detection can be divided into ECG measurement position and severity of noise or artifacts. However, when using shirt-type wearable devices for ECG recording, the shape of a QRS complex might be apparently different due to contaminated noise or artifacts, even if the positions of the measurement electrodes have not significantly changed.

Hence, in this experiment, we target an artificially generated pseudo ECG by mixing noise and artifacts into the ECG while fixing the chest lead, as in our previous study [9]. Our previous work indicated that measurement faults occurring when using shirt-type ECG devices may be similar to the combination of three irregularities defined in the MIT-BIH Noise Stress Test Database (NSTDB) [35], [36]. We therefore regarded the MIT-BIH Arrhythmia Database (MITDB) [37], [38] as the ECG and NSTDB [35], [36] as irregular waves (noise or artifacts) and mixed them as

$$targetECG = ECG_{MITDB} + k \times irregularWave_{NSTDB} \quad (16)$$

where k means a real number.

Assuming a certain measurement fault that causes noise or artifacts to suddenly occur in daily life due to body movements or physical impact on the measurement electrodes, we replaced a part of the ECG with pseudo ECGs obtained by (16). As an initial evaluation, the total data length of this experiment was set to the first 60 s, and the replaced part was set to 30 to 40 s of the target data. Because the sampling rate of MITDB and NSTDB is the same (360 Hz), we synchronized

ECG_{MITDB} and $irreggularWave_{NSTDB}$ with the same sample number, and set k in (16) to 3, whose generated pseudo ECG was similar to that recorded by wearable ECG devices utilizing dry electrodes based on the previous study [9].

As ECG_{MITDB}, we utilized V5 derived ECGs of ID no. 100, whose chest lead makes it easy to detect QRS complexes in the normal measurement state. In the experiment target (i.e., the first 60 s), we confirmed that there were very few measurement faults and we were able to clearly observe QRS complexes.

As $irreggularWave_{NSTDB}$, we used the data provided as "noise 1" of baseline wander (BW), electrode motion artifact (EM), and muscle artifact (MA) in NSTDB [36]. Because NSTDB provides three types of noise or artifacts, combining them all yields eight conditions: (i) RAW (i.e., without any noise or artifacts), (ii) BW, (iii) EM, (iv) MA, (v) BW+EM, (vi) BW+MA, (vii) EM+MA, and (viii) BW+EM+MA. When mixing two or more irregular waves, we set the added value of irregular waves as $irreggularWave_{NSTDB}$. Fig. 5 shows all the target pseudo ECGs that were created in accordance with the aforementioned procedures.

Because the complex wavelet utilized in the proposed method assumes the ECG data sampled at 200 Hz, all the target ECG data were down sampled after the additive synthesis using (16). To detect QRS complexes with the original performance of each QRS complex detection algorithm, no pre-processing was used.

2) REFERENCE QRS COMPLEXES

Because measurement faults themselves only cause apparent changes, without changing the heart activity itself, QRS complexes observed during measurement faults (i.e., during 30 to 40 s) were constant. We therefore used "annotations" of MITDB ID no. 100 in the first 60 s as the reference of QRS complexes in all the patterns from (i) to (viii).

In the first 60 s, 75 points are annotated. On the basis of the guidelines of PhysioBank annotation [39], we did not use the very first annotation "+" observed at 0.050 s because it was not related to actual QRS complexes. However, we used the annotation "A" observed at 5.678 s as a reference because premature atrial contraction represented as "A" does not cause changes in the shape of a QRS complex based on its clinical definition [28]. Overall, 74 points from 0.214 to 59.508 s were used as the references of QRS complexes.

For the appropriate validation of F1 score (described in detail below), we moved each annotation on the local maxima of the corresponding R wave, as we confirmed that each annotation was not necessarily affixed to the R wave but rather around the S wave.

3) PERFORMANCE EVALUATION MEASURES FOR QRS COMPLEX DETECTION

In this experiment, we first calculated precision and recall and then calculated F1 score in accordance with the following

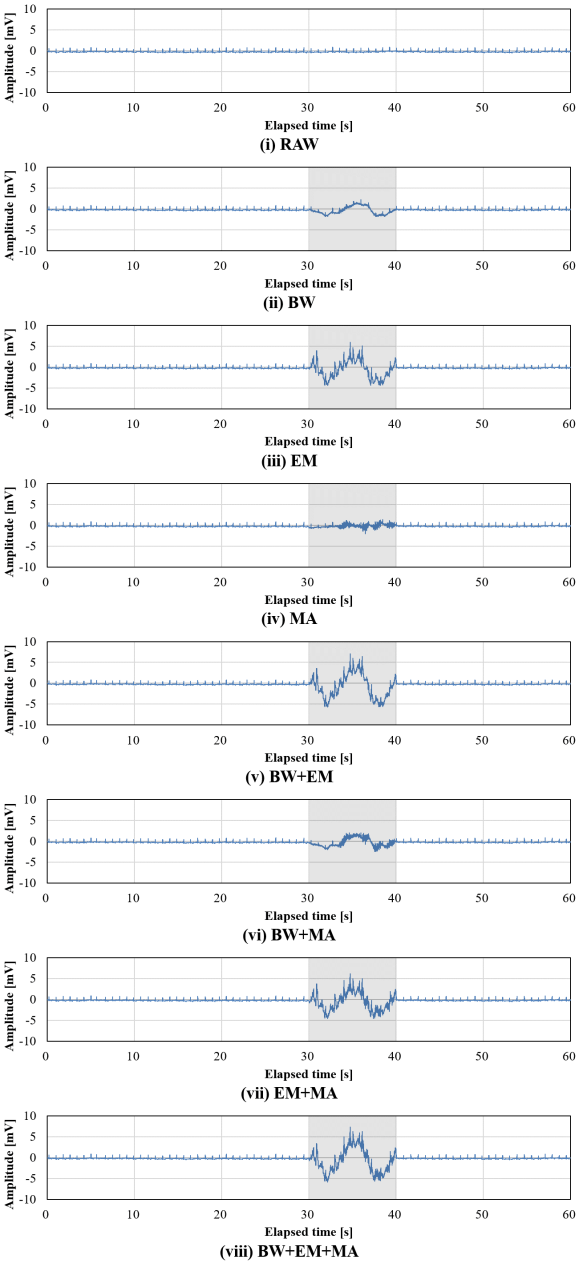

FIGURE 5. Target pseudo ECGs in experiment 1. Black shaded area in each ECG indicates the period when the target pseudo ECG undergoes measurement faults created using (16).

definitional equations [31], [32] (see Appendix A for details).

$$Precision = \frac{TP}{TP + FP} \tag{17}$$

$$Recall = \frac{TP}{TP + FN} \tag{18}$$

$$F1\ score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{19}$$

In these definitional equations, both precision and recall are theoretically influenced by the amount of true positives (TPs), so TPs must be counted on the basis of whether the

detected points are exactly related to the corresponding reference QRS complexes. To obtain TPs that are as accurate as possible in this sense, we additionally considered two perspectives: the discrimination of misdetected QRS complexes that are not related to QRS complexes, at least in terms of their observation time, and the time difference between the actual observation time of the reference QRS complex and the estimated observation time of the possible QRS complex obtained by the QRS complex detection algorithm. Regarding the first perspective, simply counting up the detected points against the number of reference QRS complexes might result in overvaluation of the QRS complex detection performance because each QRS complex detection algorithm possibly detects points irrelevant to the true QRS complexes, especially when processing ECGs with artifacts whose frequency characteristics are quite similar to those of the true QRS complexes. As for the second perspective, counting up based only on the actual observation time of the reference QRS complexes without considering the time difference between the reference QRS complex and the estimated point might lead to undervaluation, as the timing of possible QRS complex detection depends on each QRS complex detection algorithm and not necessarily on the exact observation time of the local maxima of the R wave.

To obtain all the target measures as accurately as possible while suppressing overvaluation or undervaluation, we only counted a detected point as TP when it was observed within 0.10 s from the actual observation time of the corresponding reference QRS complex. Here, 0.10 s is the normal duration of one QRS complex among healthy subjects based on its clinical definition [28]. Because we moved each annotation to the local maxima of the corresponding R wave (as described above), only the detected point observed before and after 0.10 s from the local maxima of the R wave became the counting-up target for TP.

4) PERFORMANCE EVALUATION OF RRI CALCULATION

To evaluate the performance of each QRS complex detection algorithm from the perspective of the effect on the subsequent RRI calculation as the initial step of HRV analysis, we also evaluated the calculated RRI by its tachogram. To determine the tolerance of the QRS complex detection performance to measurement faults, we conducted a comparative evaluation of the tachogram obtained by each QRS complex detection algorithm while focusing on three phases in this experiment: namely, the first 30 s without measurement faults, 30–40 s during measurement faults, and after 40 s back to recording ECG without measurement faults.

5) RESULTS

Fig. 6 shows the cross-sectional results of F1 score, precision, and recall obtained by each method, and Table. 1 shows the individual results in each condition obtained by each method. The overall performance in QRS complex detection based on the averaged F1 score was the best in method (a), followed by method (b) and then (c). T-tests using Bonferroni correction showed that there were significant differences in the F1 score between methods (a) and (b) ($p = 0.044$). Regardless of the method, all performed better in conditions (i) RAW and (ii) BW, and worse in (vii) EM+MA and (viii) BW+EM+MA. This indicates that all three methods were tolerant of the ''noise'' represented as BW, where QRS complexes were easily observable even by sight, but vulnerable to ''artifacts,'' especially when QRS complexes were comparatively harder to observe.

Fig. 7 shows the results of the RRIs calculated from the QRS complexes obtained by each method. Although the average F1 score for each method was over 0.85, the detection performance of QRS complexes was apparently quite different. Both methods (a) and (b), which utilized a single complex wavelet, returned to the original performance right after the end of measurement faults regardless of the conditions, whereas method (c) could not do that in conditions (vii) EM+MA or (viii) BW+EM+MA. Specifically, method (c) in these conditions could not detect QRS complexes for approximately 5 seconds after the end of measurement faults. Comparing methods (a) and (b) from 30 to 40 s (i.e., during measurement faults), method (a) was able to obtain a comparatively closer value to the reference RRIs.

The above results demonstrate that the proposed method (a) performed the best among the three target methods in terms of both QRS complex detection performance and RRI calculation based on the detected QRS complexes.

B. EXPERIMENT 2: EVALUATION OF QRS COMPLEX DETECTION PERFORMANCE ON ACTUAL ECG DATA RECORDED BY SHIRT-TYPE WEARABLE ECG DEVICE ASSUMING DAILY LIFE ACTIVITIES

In this experiment, we assumed a situation we often encounter when recording ECGs by shirt-type wearable ECG devices in a daily life environment. Specifically, we evaluated the performance of the proposed method targeting actual ECGs recorded during exercise by a shirt-type wearable ECG device.

Because shirt-type wearable ECG devices possibly lead to measurement faults in which we cannot confirm any QRS complexes by sight, we conducted a comparative evaluation using two independent wearable ECG devices at the same time: a shirt-type device for the performance evaluation of each QRS complex detection algorithm and a patch-type device with medical approval for the reference. The subject first wore the patch-type device for the reference and then the shirt-type device for the performance evaluation. After testing both devices, we confirmed that each device was able to record ECGs without interfering with the ECG recording by the other device.

To enable a fair comparative evaluation using these two independent devices, we conducted a pre-evaluation of each one: a pre-evaluation of the reference QRS complexes obtained by the patch-type wearable ECG device with medical approval, and a pre-evaluation of analysis target ECGs

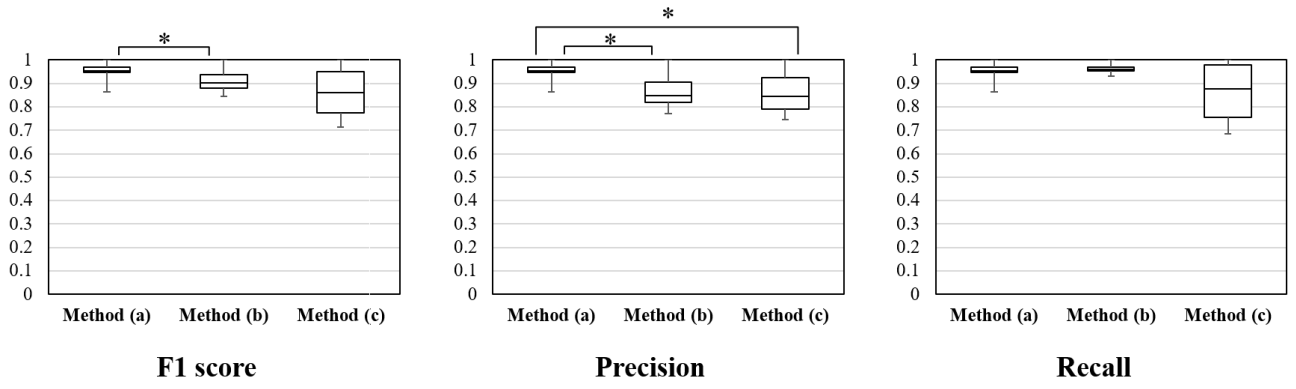

FIGURE 6. Cross-sectional results of F1 score, precision, and recall obtained by each method in experiment 1. T-tests using Bonferroni correction showed there were significant differences in the F1 score between methods (a) and (b) ($p = 0.044$), in precision between methods (a) and (b) ($p = 0.017$) as well as between methods (a) and (c) ($p = 0.031$).

TABLE 1. Comparison of QRS complex detection performance in experiment 1.

	F1 score			Precision			Recall		
	Method (a)	Method (b)	Method (c)	Method (a)	Method (b)	Method (c)	Method (a)	Method (b)	Method (c)
(i) RAW	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
(ii) BW	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
(iii) EM	0.959	0.897	0.800	0.959	0.843	0.806	0.959	0.959	0.795
(iv) MA	0.945	0.915	0.934	0.945	0.875	0.899	0.945	0.959	0.973
(v) BW+EM	0.959	0.892	0.792	0.959	0.833	0.803	0.959	0.959	0.781
(vi) BW+MA	0.945	0.903	0.921	0.945	0.854	0.886	0.945	0.959	0.959
(vii) EM+MA	0.863	0.845	0.714	0.863	0.773	0.746	0.863	0.932	0.685
(viii) BW+EM+MA	0.945	0.845	0.719	0.945	0.773	0.758	0.945	0.932	0.685
Average	0.952	0.912	0.860	0.952	0.869	0.862	0.952	0.962	0.860
Standard deviation	0.040	0.056	0.111	0.040	0.083	0.094	0.040	0.024	0.129

for the QRS complex detection algorithm recorded by the shirt-type wearable ECG device.

1) TARGET EXERCISE

Assuming the ECGs recorded by shirt-type wearable ECG devices during daily life activities, this experiment targeted the ECGs recorded while the subject performed "radio exercise no. 1 [40]" (see Appendix B for more detail). This exercise was originally designed to help people improve their physical fitness, so it involves exercises for all parts of the body, including jumping and twisting/stretching/bending of the trunk. Because exercises involving trunk movements may cause slippage or movement of the measurement electrode embedded inside the shirt, which results in recording ECGs contaminated with noise and artifacts, we expected we could validate the potential performance of each QRS complex detection algorithm during daily life activities by targeting the ECGs recorded during this exercise.

This exercise consists of 13 short workouts (shown in Table. 2) that are easily performed by ordinary people from children to the elderly. Each workout is set to music with short commands on the next workout, and the entire set takes approximately 3 min from start to finish [41]. In accordance with the approximate duration of each workout in the video (see Table. 2), we targeted the ECG recorded for 200 s from the start of the designated music. The first 150 s of the video in increments of 15 s correspond to the first to tenth workouts, the next 10 s correspond to the 11th workout, the next 15 s correspond to the 12th workout, and the final 25 s correspond to the 13th workout.

Before the experiment, all participants watched the video and practiced the movements. In case a participant forgot any of the movements, we also played the exact same video during the experiment.

2) REFERENCE QRS COMPLEXES AND CORRESPONDING ECG DATA RECORDING

The reference QRS complexes in this experiment were separately recorded by a Holter ECG monitoring device with medical approval. An overview of the device is shown in Fig. 8.

The ECG data were recorded by the combination of a wearable ECG device with built-in analog-to-digital (A/D) converter (Cardy 303 pico+, SUZUKEN CO., LTD., Aichi,

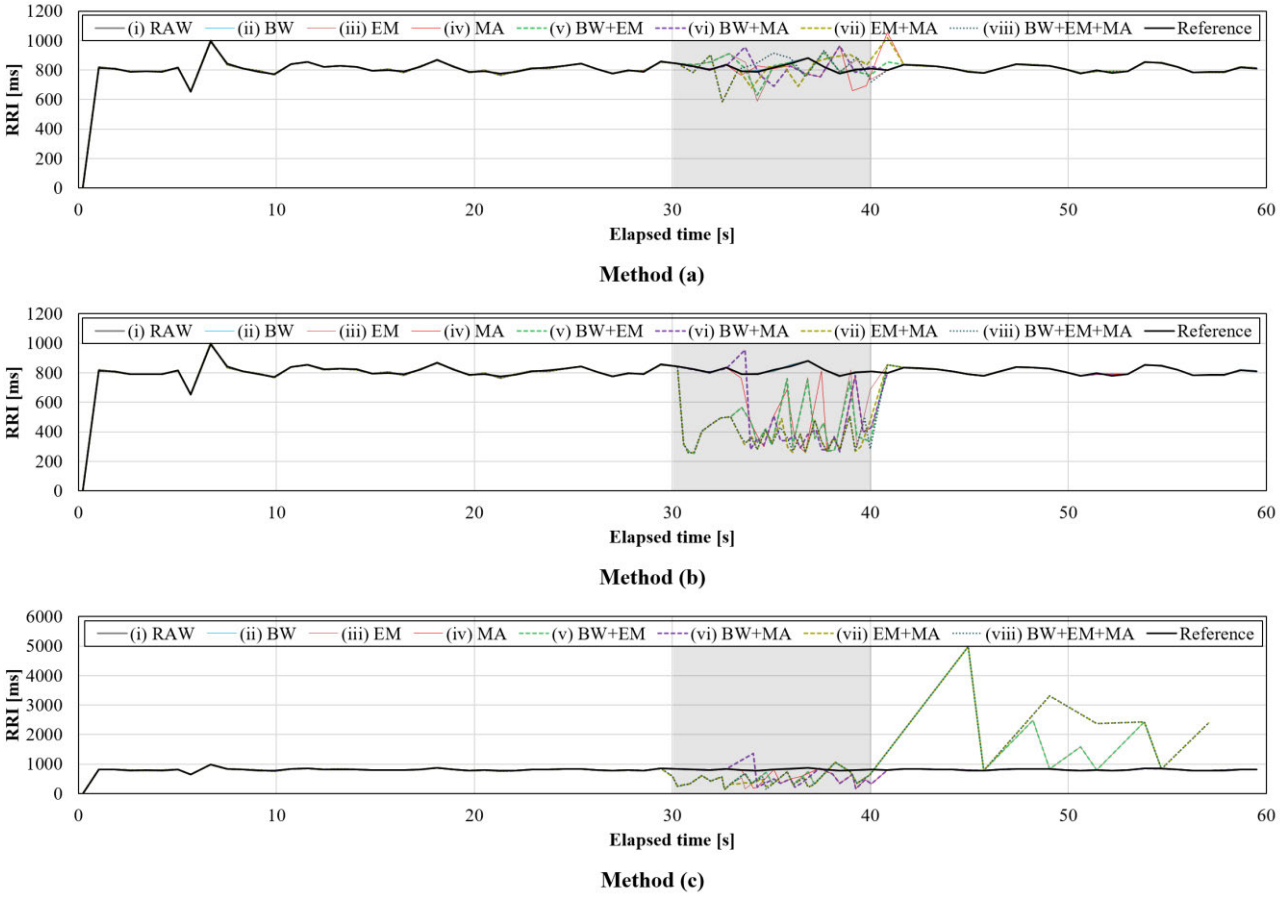

FIGURE 7. RRIs calculated from QRS complexes obtained by each method. Black shaded area indicates the period when the target pseudo ECG undergoes measurement faults created using (16). Note that the value range of the longitudinal axis in method (c) was different from that in methods (a) and (b), because method (c) could not detect any QRS complexes for a while in several conditions.

TABLE 2. 13 Workouts comprising radio exercise no. 1 and its duration.

No.	Workout	Duration [s]
1	Stretching the body	15.83*
2	Swinging the arms and bending/stretching the legs	14.38
3	Rotating the arms	14.28
4	Spreading the chest	15.13
5	Side bending of the body	14.32
6	Bending of the body back and forth	13.99
7	Body twisting	13.26
8	Stretching the arms up and down	13.74
9	Bending the body diagonally downwards and spreading the chest	15.80
10	Rotating the body	15.49
11	Jumping with both legs	9.41
12	Swinging the arms and bending/stretching the legs (the same as workout no.2)	14.70
13	Deep breaths	21.27

The authors measured the time of each workout by a stopwatch.
*Including 8.58 s for "pause"

Japan) and its dedicated disposable water-resistant patch-type electrode (Kenz Cardyrode-P [waterproof], SUZUKEN CO., LTD., Aichi, Japan). The size of the wearable ECG device was 28 mm [width] × 42 mm [height] × 9 mm [depth]

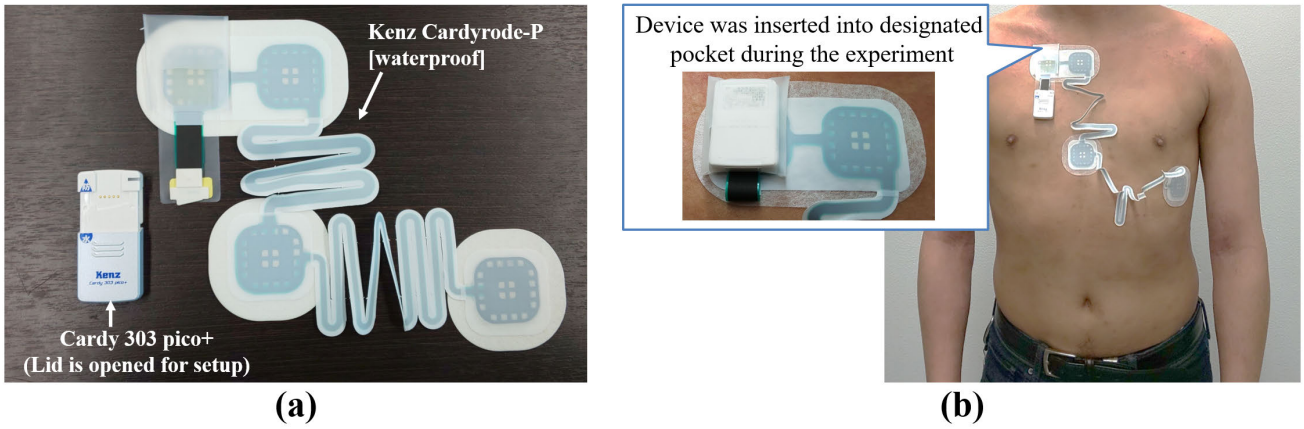

FIGURE 8. Overview of the wearable Holter ECG monitoring device for obtaining reference QRS complexes. (a) Complete set of the wearable device (Cardy 303 pico+ and Kenz Cardyrode-P [waterproof], SUZUKEN CO., LTD., Aichi, Japan). (b) Installation example.

TABLE 3. ECG leads of the wearable Holter ecg monitoring device and its channel information according to "instructions for use" of Cardy 303 pico+.

Channel	Lead	Anode (+)		Cathode (-)		Ground (GND)	
CH1	CM5	3	On the rib near V5	1	Upper end of breastbone	4	Position of the electrode where naturally extended from the position of electrode 1 (upper end of breastbone)
CH2	NASA	2	Lower end of breastbone	1	Upper end of breastbone		
CH3	Auxiliary lead	3	On the rib near V5	2	Lower end of breastbone		

and its weight was 13 g. The dedicated disposable water-resistant patch-type electrode consisted of three integrally formed patches and lead wires. The size of the one patch with two electrodes was 102 mm [width] × 57 mm [height], whereas the two patches with one electrode were 57 mm square. Each independent patch (i.e., electrode unit) was connected with a lead wire. The weight of this integrated patch was 8 g, so that the total weight of a complete set of this device worn by users was 21 g. Thanks to the specially designed electrodes, we could record three-channel ECGs simultaneously (i.e., NASA, CM5, and an auxiliary lead, as shown in Table. 3 and Fig. 9) at a sampling rate of 125 Hz. This combinative ECG recording enables the suppression of measurement faults compared to other devices that only record single-channel ECGs.

After each participant finished the experiment, the recorded ECGs were transferred from the wearable ECG device (i.e., Cardy 303 pico+) to a personal computer by using its dedicated wired transfer unit (Cardy Transfer Unit 03, SUZUKEN CO., LTD., Aichi, Japan) and then analyzed using dedicated software (Kenz Cardy Analyzer Lite, SUZUKEN CO., LTD., Aichi, Japan). Both the wired transfer unit and the software are registered as medical instruments associated with the aforementioned ECG device.

The dedicated software is only capable of outputting the list of recorded RRIs without raw ECG data. Because RRIs are calculated right after the detection of a QRS

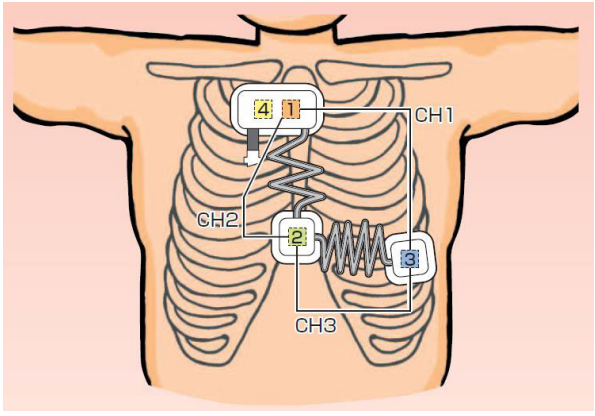

FIGURE 9. ECG leads of the wearable Holter ECG monitoring device. This figure is cited from "Instructions for use" of Cardy 303 pico+.

complex in theory, we regarded the observation time of each RRI as the reference of the QRS complex observation time.

3) ANALYSIS TARGET ECG DATA FOR QRS COMPLEX DETECTION ALGORITHM AND ITS RECORDING

The analysis target ECG data in this experiment were recorded using a commercial shirt-type wearable ECG device. Fig. 10 shows the device overview. It consists of a specially designed shirt with the embedded measurement

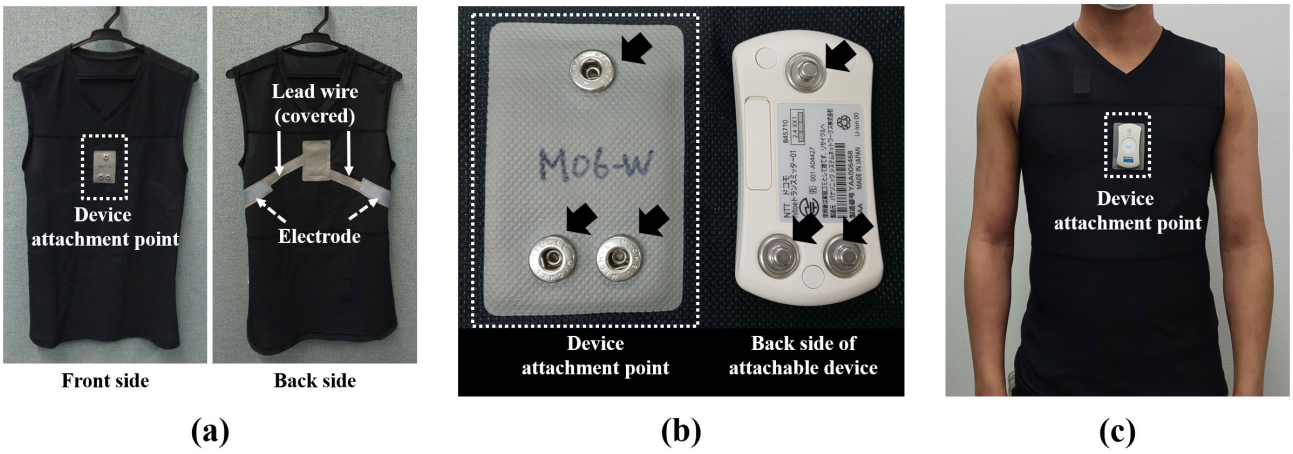

FIGURE 10. Overview of the shirt-type wearable ECG device for recording analysis target ECG. (a) Shirt with embedded electrodes and lead wire (Toray Industries, Inc., Tokyo, Japan). (b) Device attachment point on the front side of the shirt and the dedicated attachable hardware for ECG recording (hitoe® transmitter 01, NTT DOCOMO, INC., Tokyo, Japan). (c) Installation example. The black arrows in (b) indicate the snap buttons for attachment.

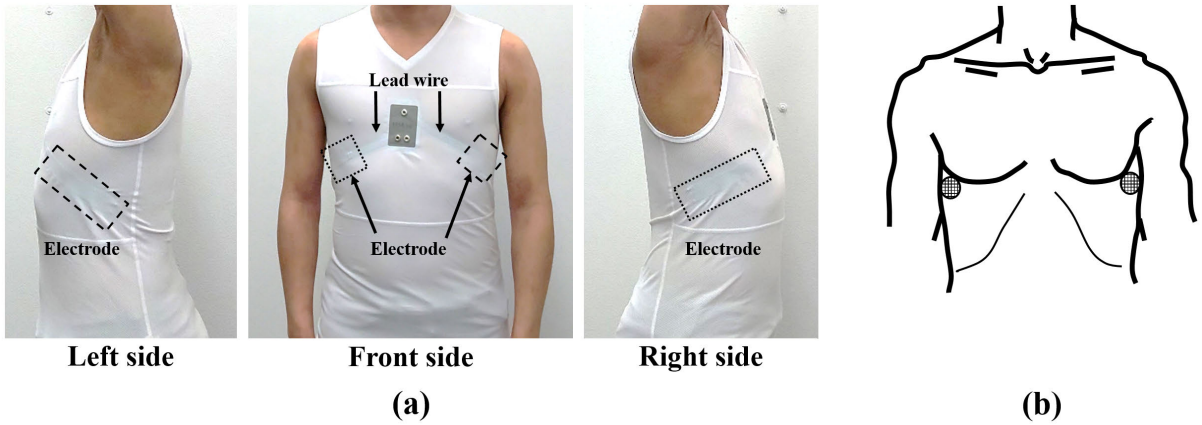

FIGURE 11. ECG leads of the shirt-type wearable ECG device. (a) Electrode position on the shirt. (b) Bipolar chest lead CC5. For easy understanding of the electrode position, (a) shows the white colored version of the same size shirt in Fig. 10. In (b), circles with a lattice pattern indicate the position of measurement electrodes in bipolar chest lead CC5. Note that the image in (b) was created by the authors based on [43].

electrode and lead wire (Toray Industries, Inc., Tokyo, Japan) and an attachable wearable ECG device (hitoe® transmitter 01, NTT DOCOMO, INC., Tokyo, Japan). The embedded measurement electrodes were made of the functional material hitoe® (Toray Industries, Inc., Tokyo, Japan), which is electroconductive textile fabric made of nano-fiber yarn (fiber diameter 700 nm; polyester) coated with a PEDOT-PSS polymer thermobonding composition [6], [42] sized 130 mm $\times$ 26 mm. The measurement electrodes made of hitoe® were installed in the inside of the shirt with slight compression, which enabled the measurement electrodes to be snugly placed on the skin surface. The parts other than the aforementioned measurement electrode (e.g., lead wire and the back side of the device attachment point) were coated with a non-conductive material (Fig. 10(a)), so that the attachable wearable ECG device only records the signal from measurement electrodes via snap buttons for attachment.

Regarding the size of the shirt, five sizes (SS, S, M, L, and LL) were available depending on the physique of the subject, and its weight was 112.0 $\pm$ 11.64 g (min: 94 g in SS size, max: 130 g in LL size). The combination of the specially designed shirt with embedded measurement electrodes and an attachable wearable ECG device was capable of recording a single-channel ECG whose measurement lead is similar to a bipolar chest lead CC5 [43] (Fig. 11).

Before each experiment, we checked the size of the shirt to make sure the subject could wear it without having the electrode move away from the skin or bend.

The ECG data were recorded through a wearable ECG device (hitoe® transmitter 01, NTT DOCOMO, INC., Tokyo, Japan) comprising a built-in A/D converter and a wireless transmitter. The size of the wearable ECG device was 37 mm [width] $\times$ 70 mm [height] $\times$ 12 mm (thickest) [depth] and its weight was 23 g. The ECG data recorded at a

sampling rate of 200 Hz were processed through this built-in A/D converter and wirelessly transmitted to a smartphone (Xperia^{TM} X performance [SO-04H], Sony Mobile Communications Inc., Tokyo, Japan) via a dedicated application in real-time.

After finishing all the experiments, all ECG data were transferred to a personal computer, and the designated QRS complex detection algorithm was applied for off-line processing.

4) PERFORMANCE EVALUATION MEASURES FOR QRS COMPLEX DETECTION

All the evaluation measures (i.e., precision, recall, and F1 score) were calculated in accordance with (17)–(19) described in III.A.3. To obtain TPs that are as accurate as possible, it is important to consider the synchronizing method of the two different devices as well as the suppression method of their difference in the sampling rate (i.e., 125 Hz in the reference and 200 Hz in the target ECG data) in addition to the aforementioned two perspectives described in III.A.3.

To unify the time counting systems of the two independent devices, we used the time elapsed since the observation time of the first QRS complex. Because of the difference in the sampling rate of each device, we could not use their output time as is (i.e., 24-hour clock with the calendar date). Furthermore, the wearable ECG device for recording the reference utilized stand-alone operation and was only able to output the list of RRIs without raw ECG data, so we could not synchronize two independent devices at the level of analog signal. Because the first 9 s of the target exercise was ''pause,'' without any physical movements where we were able to clearly confirm QRS complexes in both devices regardless of the subject, we synchronized the two devices by the observation time of the first QRS complex while setting that time as the starting point of the elapsed time (i.e., 0 s).

Regarding the sampling rate, we assumed that its effect on the observation time of a QRS complex was comparatively smaller (i.e., 0.003 s) than the time difference between the actual observation time of the reference QRS complex and the estimated observation time of the possible QRS complex obtained by the QRS complex detection algorithm. We therefore aimed to resolve this difference along with the two perspectives mentioned earlier.

To obtain all the target measures as accurately as possible while considering the remaining perspectives, we only counted a detected point as TP when it was observed within 0.10 s from the actual observation time of the corresponding reference QRS complex, the same as in experiment 1. We regarded the update timing of each reference RRI as the actual observation time of a QRS complex. Because the dedicated software (Kenz Cardy Analyzer Lite, SUZUKEN CO., LTD., Aichi, Japan) can detect QRS complexes within the time it takes to observe them, we presumed that the time difference between the actual QRS complex and the observation time of RRI would be sufficiently small for evaluating TPs.

5) PERFORMANCE EVALUATION OF RRI CALCULATION

We also evaluated the calculated RRI by its tachogram in this experiment. To determine the tolerance of the QRS complex detection performance to measurement faults, we conducted a comparative evaluation of the tachogram obtained by each QRS complex detection algorithm and comparing it with the recorded ECG, as the QRS detection performance depends on the quality of the recorded ECG itself.

6) PRE-EVALUATION OF REFERENCE QRS COMPLEXES

Fig. 12 shows the RRIs obtained from each subject by using a Holter ECG device.

As seen in Fig. 12(i), there were no misdetections or overlook of RRIs in the case of subject 1. Regarding subjects 2 (Fig. 12(ii)) and 3 (Fig. 12(iii)), on the other hand, some RRI calculation failures were confirmed: according to its annotation, the device failed to obtain RRIs observed at 170.28 s and 170.56 s for subject 2, and at 85.856 s and 86.144 s for subject 3. We therefore excluded these four RRIs and used the rest of them as the reference QRS complexes. In total, the following points were set as the reference QRS complexes in each subject: 274 points for subject 1, 332 points for subject 2, and 321 points for subject 3.

7) PRE-EVALUATION OF ANALYSIS TARGET ECGS FOR QRS COMPLEX DETECTION ALGORITHM

Fig. 13 shows the analysis target ECGs for the QRS complex detection algorithm recorded by using the aforementioned shirt-type wearable ECG device during the target exercise.

As seen in Fig. 13(iii), the ECG obtained from subject 3 was comparatively clear regardless of the movements. Regarding subjects 1 (Fig. 13(i)) and 2 (Fig. 13(ii)), on the other hand, several parts of the recorded ECG were contaminated with noise or artifacts. As a whole, the ECG obtained from subject 1 was in poor condition compared to the one obtained from subject 2. Focusing on each workout in the target exercise, the ECGs recorded during workouts 3 to 10 were relatively worse than those of the other workouts. Because workouts 3 to 10 involved twisting/stretching/bending of the trunk, which may cause slippage or movement of the measurement electrode embedded inside the shirt, we could validate the performance of the QRS complex detection targeting the ECGs recorded with possible measurement faults.

Taken together, we expect that the evaluation using these three ECGs as the analysis target for the QRS complex detection algorithm will help clarify the possible performance of each algorithm targeting the ECGs recorded during daily life activities, as we initially intended.

8) RESULTS

Fig. 14 shows the cross-sectional results of F1 score, precision, and recall obtained by each method, and Table. 4 shows the individual results in each condition obtained by each

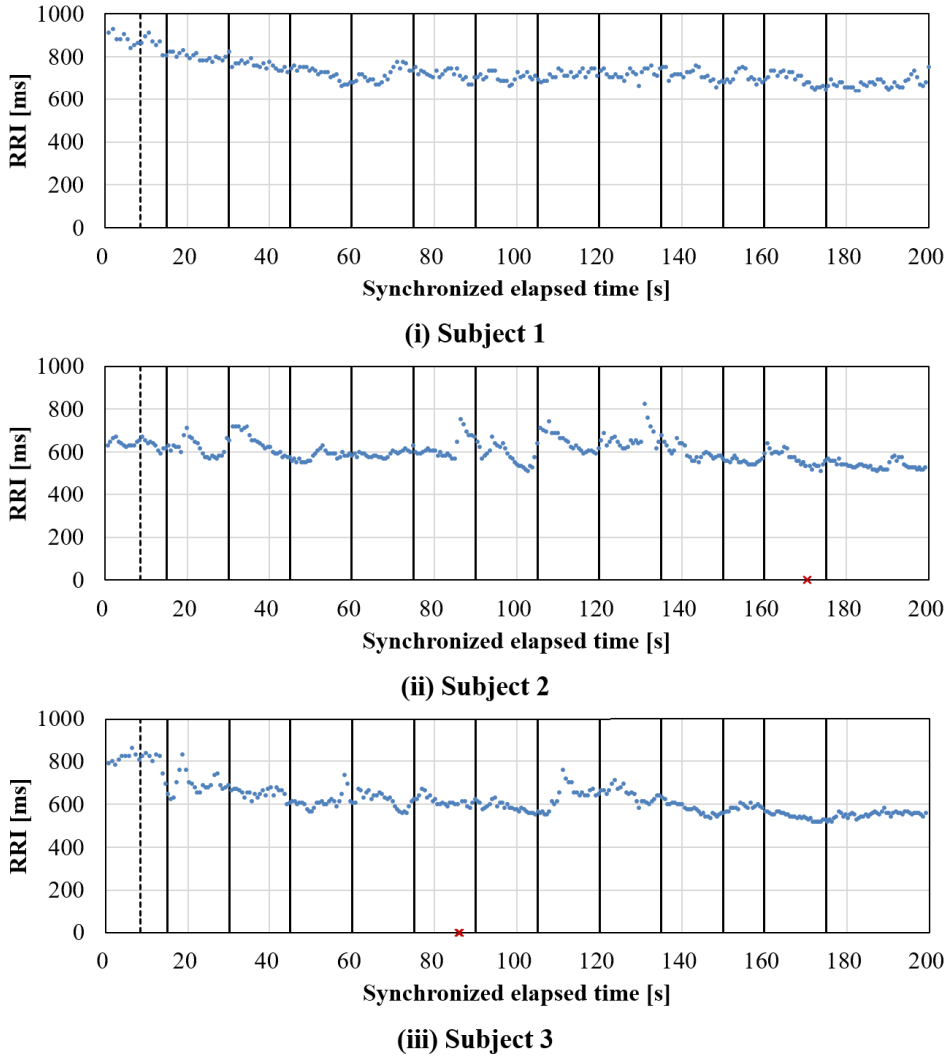

FIGURE 12. RRIs obtained from each subject by using a series of the wearable Holter ECG monitoring device. The red crosses at 170.28 s and 170.56 s in (ii) Subject 2, and at 85.856 s and 86.144 s in (iii) Subject 3, indicate erroneous RRIs according to the annotation (i.e., the RRIs excluded from the reference RRIs in experiment 2).

method. The overall performance in QRS complex detection based on the average F1 score was the best in method (a), followed by method (b) and then (c), which is the same order observed in experiment 1. T-tests using Bonferroni correction showed that there were no significant difference between any combinations. Regardless of the method, each performed better for subject 3, whose ECG was stable compared to those obtained from subjects 1 and 2. These results indicate that, while all three methods could perform well when targeting comparatively clear ECGs, their tolerances for actual measurement faults (noise or artifacts) were different.

Fig. 15 shows the results of the RRIs calculated from the QRS complexes obtained by each method. As we confirmed in the aforementioned F1 score, all methods performed well for subject 3, whose ECG was stable compared to those obtained from subjects 1 and 2. Regarding the other subjects, each method performed differently. Method (a) was able to obtain RRIs that were apparently close to the reference RRIs regardless of the subject, whereas methods (b) and (c) failed to obtain applicable RRIs, especially when processing ECGs with measurement faults (i.e., the ECGs obtained from subjects 1 and 2). This tendency was worse in method (c): when processing subject 1, it could not match to the original performance and obtained 7285 ms as one RRI at worst.

Considering these results, the combination of preprocessing by complex wavelet and the proposed peak detection, which was performed as method (a) in this experiment, was effective to acquire tolerance to measurement faults (e.g., during/after measurement faults).

IV. DISCUSSION

Assuming that the QRS complex detection algorithm will be used for ECGs recorded by a wearable ECG device in the daily life environment as the initial step of HRV analysis,

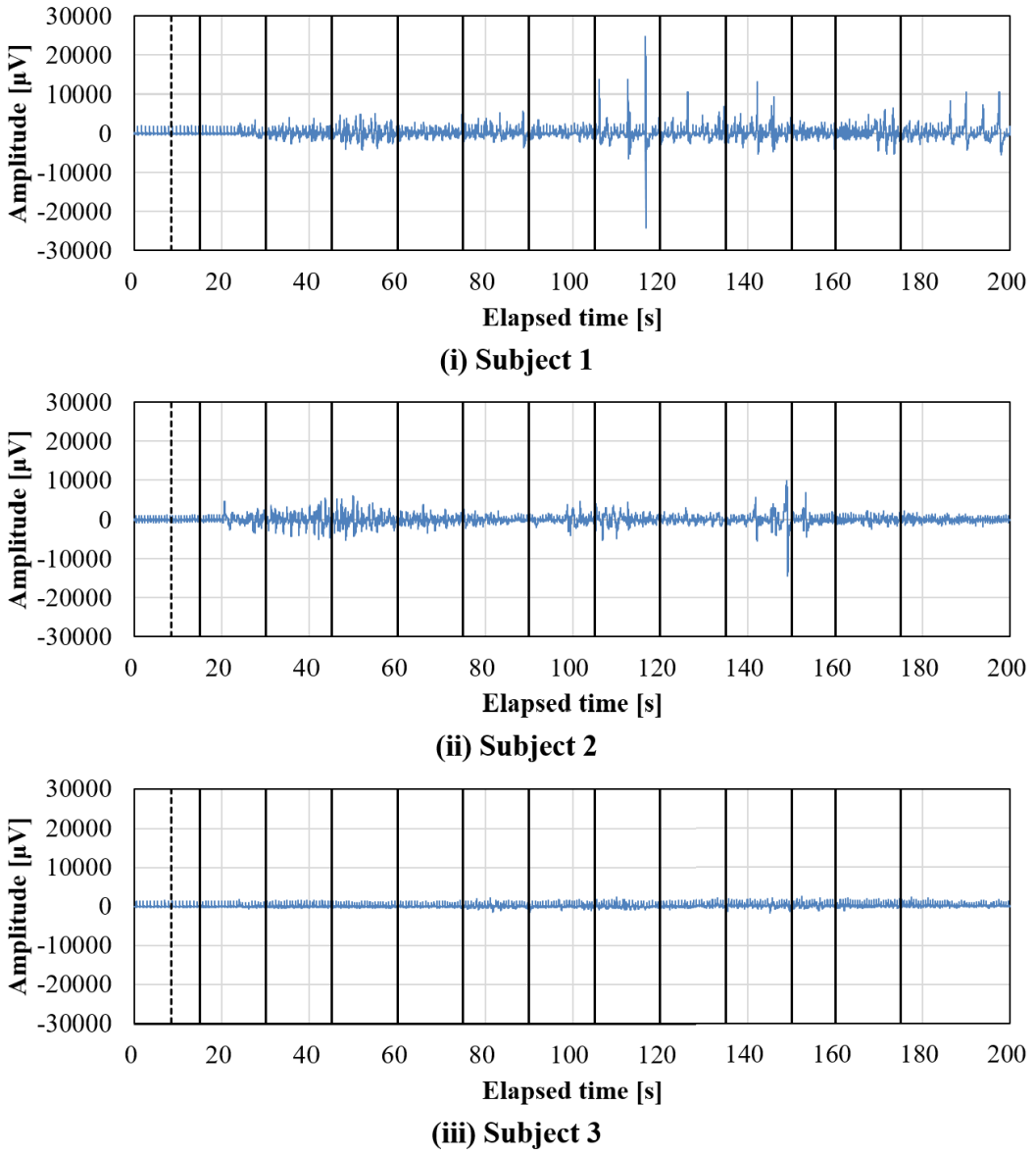

FIGURE 13. Analysis target ECGs for QRS complex detection algorithm recorded by using the shirt-type wearable ECG device during target exercise.

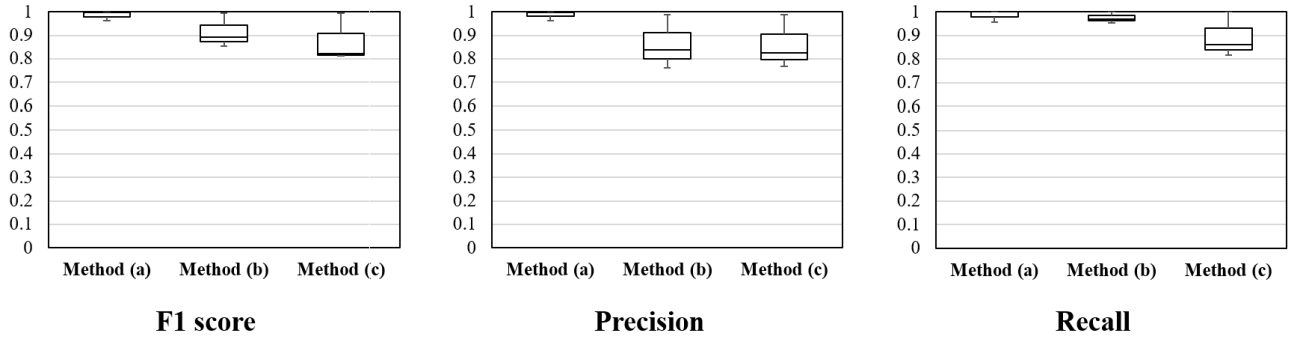

FIGURE 14. Cross-sectional results of F1 score, precision, and recall obtained by each method in experiment 2. T-tests using Bonferroni correction showed there were no significant differences between any combinations.

TABLE 4. Comparison of QRS complex detection performance in experiment 2.

	F1 score			Precision			Recall		
	Method (a)	Method (b)	Method (c)	Method (a)	Method (b)	Method (c)	Method (a)	Method (b)	Method (c)
Subject 1	1.000	0.853	0.821	1.000	0.761	0.824	1.000	0.971	0.818
Subject 2	0.961	0.892	0.813	0.964	0.837	0.769	0.958	0.955	0.862
Subject 3	0.998	0.994	0.994	0.997	0.988	0.988	1.000	0.971	1.000
Average	0.986	0.913	0.876	0.987	0.862	0.860	0.986	0.975	0.893
Standard deviation	0.018	0.059	0.083	0.016	0.094	0.093	0.020	0.019	0.077

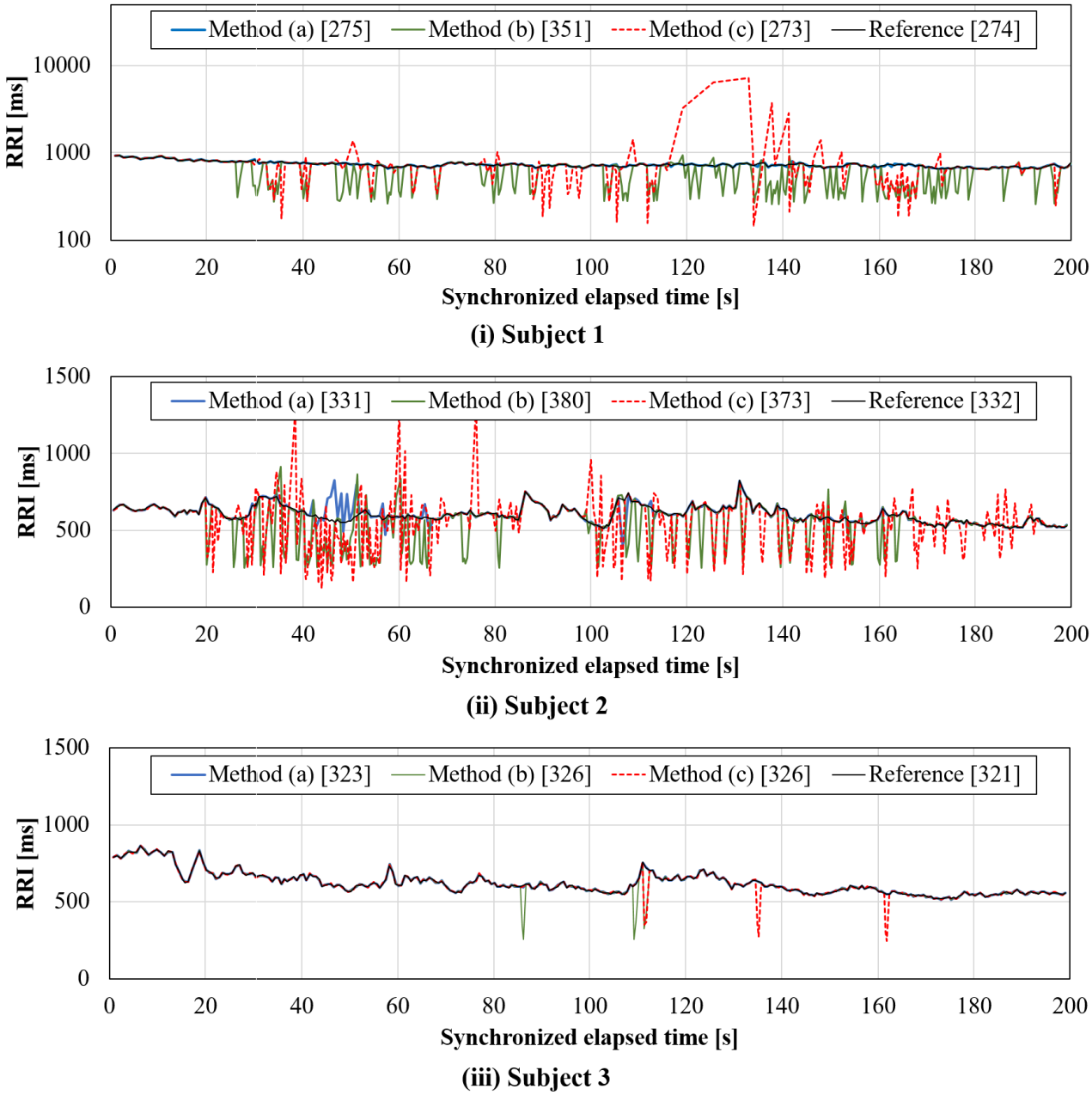

FIGURE 15. Results of RRIs calculated from the QRS complexes obtained by each method. Figures in brackets indicate the amount of the estimated RRIs obtained by each method. Note that the value range of the longitudinal axis in (i) subject 1 utilized a logarithm whereas the others utilized a real number, as method (c) for subject 1 could not detect any QRS complexes for a while and therefore calculated prolonged RRIs due to the overlooked QRS complexes.

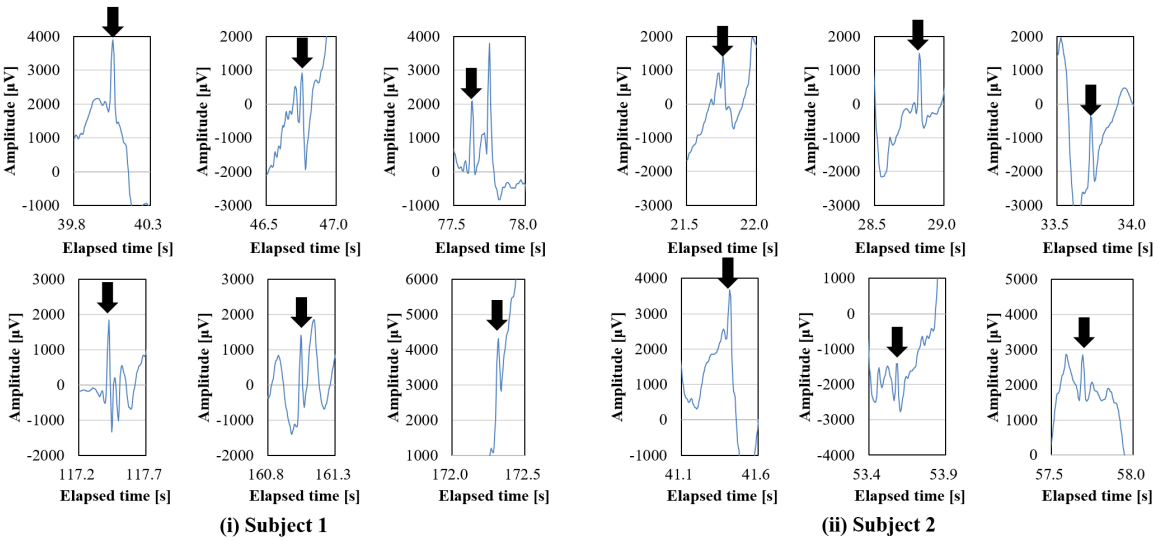

FIGURE 16. Examples of QRS complexes detected by the proposed method that the benchmark method (Pan-Tompkins algorithm [11], method (c)) could not detect. Black arrows indicate the positions of QRS complexes. All these QRS complexes were confirmed to correspond to the reference RRIs obtained by a Holter wearable ECG device, in which their annotation indicated "normal" beat, so all morphological changes in these QRS complexes were considered apparent changes due to measurement faults (e.g., noise or artifact).

the following two perspectives should be satisfied: (α) tolerable performance in detecting QRS complexes while suppressing overlook and misdetection that can be accomplished with only a small dataset (or no dataset at best) and (β) operability on limited computing resources in real-time or semi-real-time (i.e., with slight delay). Because the currently available wearable ECG devices transfer the recorded ECG as batch processing after finishing its recording (e.g., Holter ECG devices such as the Cardy 303 pico+ used in this experiment) or as real-time/semi-real-time processing using the combination of a dedicated wireless transmitter (e.g., the hitoe® transmitter 01 used in this experiment) and a commercial smartphone, QRS complex detection should be performed either in the built-in microcomputer in the transmitter or in the mobile environment represented as a smartphone, neither of which can utilize a sufficient dataset for processing QRS complex detection.

To satisfy condition (α) in the design stage of the algorithm, we developed a two-stage QRS complex detection algorithm comprising single complex wavelet filtering in the preprocessing stage and morphology-based peak selection in the decision stage. To line up candidates of the QRS complexes while suppressing overlook regardless of the change in the shape of QRS complexes without preparing any learning dataset in advance, we utilized single complex wavelet filtering in the preprocessing stage that can follow the change in shape of QRS complexes simply by changing phases (Fig. 4). Through experiments 1 and 2 targeting ECGs assuming daily life activities, we confirmed that the proposed single complex wavelet (methods (a) and (b)) was able to obtain the QRS complexes corresponding to the reference ones with more accuracy than the algorithm utilizing different preprocessing (method (c)). Fig. 16 shows examples of the QRS complexes detected by the proposed method (method (a)) from the actual ECG data in experiment 2, which the benchmark method (method (c)) could not detect. Although the subject wore a shirt-type wearable ECG device without any observable riding up, the shapes of the QRS complexes in the ECGs recorded during the target exercise were inconsistent due to measurement faults (noise or artifacts). Even in this situation, the proposed single complex wavelet was able to successfully line up the candidates of QRS complexes regardless of shape changes. To suppress misdetection of QRS complexes for obtaining more accurate RRIs, the proposed method utilized morphology-based peak selection in the decision stage. By making use of the phase properties in the single complex wavelet, the proposed method was better able to identify applicable QRS complexes while suppressing misdetection compared to the conventional methods. Method (b), whose decision stage was only utilized magnitude-based thresholding (i.e., "findpeaks" function in MATLAB) frequently detected unrelated points that were not QRS complexes, and method (c), whose decision stage utilized PTA built-in thresholding, sometimes failed to detect QRS complexes for a while after measurement faults occurred. Taken together, we conclude that the utilizing single complex wavelet was truly able to suppress the overlook of QRS complexes, but this alone does not necessarily contribute to suppressing the misdetection of QRS complexes; indeed, it results in increased misdetection. We therefore feel that the proposed method utilizing single complex wavelet used in conjunction with morphology-based peak selection would balance the appropriateness of the detected QRS complexes and enable tolerance to measurement faults (e.g., during/after measurement faults). The F1 scores and the calculated RRIs support this indication as

FIGURE 17. Actual run state of the dedicated mobile application utilizing the proposed method on a commercial smartphone.

well: method (a) obtained the F1 score of 0.952 ± 0.040 in experiment 1 and 0.986 ± 0.018 in experiment 2, and its calculated RRIs were also similar to the reference in both experiments.

Regarding this QRS complex detection performance evaluation, we understood that there was the possibility of undervaluation in experiment 2, where we did not have any defined annotations as in experiment 1: in theory, we cannot validate the QRS complexes detected from wearable ECG if we cannot obtain the reference QRS complexes. Focusing on subject 3, in fact, we confirmed 323 points from the wearable ECG without any drastic measurement faults in the recorded ECG, but we had only 321 reference QRS complexes. However, the number of excluded references per subject was just two points at the most thanks to the three-channel ECG recording by the wearable Holter ECG device, so that we conclude that the impact stemming from no references on these evaluation results was virtually negligible.

Although we conducted all the experiments in this paper in a commercial personal computer environment as off-line processing, we also independently investigated the potential real-time performance (i.e., in terms of condition (β)) by developing an actual mobile application. Fig. 17 shows the actual run state of the dedicated mobile application utilizing method (a) on a commercial smartphone (Xperia^{TM} X performance [SO-04H], Sony Mobile Communications Inc., Tokyo, Japan; CPU, Snapdragon 820 (MSM8996) 2.2 GHz/1.6 GHz; RAM, 3.0 GB; OS, Android 6.0.1). Triangles on the recorded ECG depict the points detected by the proposed method. We also independently confirmed that this dedicated mobile application was able to operate for 24 hours without slowdown or forced termination through a long-term stability test. Because the single wavelet designed for QRS complexes does not require much queueing time for its processing, the actual processing delay of the proposed method predominantly depends on the setting value of the queuing time for the peak selection in the decision stage, which may also affect the performance of the QRS complex detection itself. As a practical example, the experimental results in this paper indicate that the proposed method has the potential to obtain a tolerable performance within 2.4 seconds. This semi-real-time processing of QRS complex detection while recording ECG with a tolerable performance in terms of applicable RRI calculation might indirectly contribute to reducing the need for re-experiments due to the delayed confirmation of the RRI calculation performance.

It is true that improving the QRS complex detection performance itself would be an indispensable approach to improve

the accuracy of HRV features, as the RRIs used for HRV feature calculation are originally calculated from two adjacent QRS complexes in theory. With the objective of more accurate HRV analysis using wearable ECG devices in a daily life environment, however, we should also consider the subsequent processes of QRS complex detection, namely, the combination of artefact identification, RRI editing, and RRI rejection, which we can rephrase as ''RRI outlier processing,'' and HRV feature calculation itself. Although the proposed method was able to suppress overlook in QRS detection, it is actually very difficult to completely suppress the misdetection of artifacts whose frequency characteristics as well as morphology are quite similar to those of the QRS complex. Because artifacts have no relation to heart activity itself, excluding miscalculated RRIs on the basis of the reliability of the calculated RRIs as a part of RRI outlier processing would be an essential countermeasure for more accurate HRV analysis [44], [45]. Aside from this, interpolating missing RRIs caused by overlooking of a QRS complex itself, or excluding possible miscalculated RRIs, also plays an important role in maintaining the accuracy of HRV features [9], [46]. A pilot study considering the combination of several HRV analysis processes (i.e., QRS complex detection, possible misdetected RRIs exclusion, and missing RRIs interpolation) indicated that the combination of better processes would improve the accuracy of HRV features more than simply utilizing one better process [22]. We should point out here that the traditional HRV analysis flow [1] assumes only one-way processing, meaning that each process has to deal with the issues inherent in each step. In this sense, QRS complex detection that can suppress overlook and misdetection could become the basis of more accurate HRV analysis by raising the standard of RRIs in terms of amount and accuracy.

The major limitations of this study are the lack of a comprehensive comparative evaluation with QRS complex detection algorithms other than PTA, the generalizability in terms of the wearable ECG device, and the generalizability in terms of RRI estimation for purposes other than HRV analysis. Although a lot of QRS complex detection algorithms have been developed to date, in this study we conducted a comparative evaluation only with PTA as an initial validation of the proposed method. Future study aiming for a more comprehensive comparative evaluation with other QRS complex detection algorithms should be done to reveal further characteristics including the pros and cons within each algorithm especially when targeting wearable ECGs with measurement faults. Regarding the second limitation, all the above insights are based on the results of experiments using pseudo ECG or using one specific shirt-type wearable ECG device. There are many variations among the currently available wearable devices, especially in terms of the type (e.g., patch, shirt, or wrist-band), measurement electrode (e.g., arrangement or material), or even its dedicated hardware (e.g., A/D converter or built-in filter), and all these factors in addition to the environmental condition together with the subject condition under that environment could have an effect on the recorded ECGs. Aside from this, as we mentioned earlier, identical noise or artifacts will not be repeatedly observed even when the same user tries to perform the same movements, because these measurement faults are also affected by the electrode condition (e.g., wet/dry, placement position, contact area, flexion of electrode itself, or friction between electrode and skin surface), which could hardly be exactly the same as they were before. In other words, an experimental result under one combinative condition would be an example of that condition only. To develop practical healthcare services based on HRV features obtained by using wearable ECG devices, it is essential to accumulate real data and corresponding know-how on the conditions (e.g., device, environment, subject, and electrode condition) together with validation results. As for the generalizability on RRI estimation for purposes other than HRV analysis represented as arrhythmia evaluation, we should further modify the proposed method in terms of both the preprocessing stage and the decision stage. Our current preprocessing uses a single complex wavelet focusing on the morphological change of QRS complexes due to noise or artifacts, in which the duration of QRS complexes is constant, so it would not be capable of following morbid morphological changes in QRS complexes represented as a premature ventricular contraction or bundle branch block. Regarding the decision stage, on the other hand, our current peak selection assumes normal sinus rhythm and equally evaluates three properties, namely, peak magnitude, peak location, and peak morphology (phase). This means that the current peak selection would not be capable of following morbid changes in QRS complex periodicity represented as atrial fibrillation or consecutive premature atrial contraction. To extend our proposed method for the detection of QRS complexes for these clinical purposes, we should additionally consider how to follow the morbid morphological and periodicity changes while balancing the processing cost.

V. CONCLUSION

In this paper, we proposed a semi-real-time RRI estimation for wearable ECG utilizing a two-stage structure. For its preprocessing stage, the feature of the QRS complex is extracted by using a complex-valued wavelet that can adaptively fit to morphological variations of the QRS complex to improve the traceability while retaining computing resources, especially for QRS complexes undergoing morphological changes due to noise and artifacts. For its decision stage, we made use of complex-valued features and select appropriate QRS complexes in consideration of three features: peak magnitude, peak location, and peak morphology (phase). Initial evaluations showed that the QRS complex detection performance of the proposed method achieved the F1 score of 0.952 ± 0.040 when targeting pseudo ECG data created from open data assuming wearable ECGs, and of 0.986 ± 0.018 when targeting actual ECG data recorded by a shirt-type wearable ECG device during an exercise activity. Furthermore, the proposed method was able to obtain more appropriate

RRIs close to the reference RRIs by suppressing overlook or misdetection of QRS complexes. These results indicate that the proposed method has great potential for utilization in highly accurate non-clinical healthcare services based on HRV features using wearable ECGs in terms of obtaining accurate RRIs.

The major limitations of this study are the lack of a comprehensive comparative evaluation with QRS complex detection algorithms other than PTA, the generalizability in terms of wearable ECG devices, and the generalizability in terms of RRI estimation for purposes other than HRV analysis. To reveal further characteristics including the pros and cons within each algorithm, especially when targeting wearable ECGs with measurement faults, future study aiming for a more comprehensive comparative evaluation with other QRS complex detection algorithms is needed. Regarding the second limitation, all the insights were based on the results of experiments using pseudo ECGs or using one specific shirt-type wearable ECG device, so future study targeting other combinations of wearable ECGs should be done. To extend our proposed method for the detection of QRS complexes for clinical purposes such as arrhythmia evaluation, future work should additionally consider how to follow the morbid morphological and periodicity changes while balancing the processing cost.

APPENDIX

A. CALCULATION OF PRECISION, RECALL, AND F1 SCORE BASED ON CONFUSION MATRIX

This appendix introduces a brief summary of the confusion matrix and the calculation method of precision, recall, and F1 score based on it.

The confusion matrix (Table. 5) shows the combination of the prediction results and the defined label with which we can understand how accurately the target method performs [32]. Here, each class stands for the following condition.

- *True Positive (TP)*: the class for an item classified as positive whose label is positive. In this paper, it refers to the number of correctly detected QRS complexes that match the corresponding reference QRS complexes.
- *False Positive (FP)*: the class for an item classified as positive whose label is negative. In this paper, it refers to the number of misdetected QRS complexes that do not correspond to the reference QRS complexes.
- *False Negative (FN)*: the class for an item classified as negative whose label is positive. In this paper, it refers to the number of overlooked QRS complexes.
- *True Negative (TN)*: the class for an item classified as negative whose label is negative. In this paper, this class does not exist because the algorithm only targets "positive" (i.e., QRS complex) detection and there are no "negative" labels in the reference either (i.e., all reference QRS complexes are defined as "positive").

On the basis of these four data classes (or rather, three data classes comprising TP, FP, and FN), we can calculate the following measures for the performance evaluation.

- *Precision (a.k.a. Positive Prediction Value)*: the measure that quantifies how well the evaluation target (e.g., model, algorithm) avoids false positives. This measure quantifies how well the evaluation target algorithm detects QRS complexes while suppressing misdetection.

$$Precision = \frac{TP}{TP + FP}$$

- *Recall (a.k.a. Positive Rate, Sensitivity)*: the measure that quantifies how well the evaluation target (e.g., model, algorithm) avoids false negatives. This measure quantifies how well the evaluation target algorithm detects QRS complexes while suppressing overlook.

$$Recall = \frac{TP}{TP + FN}$$

- *F1 Score*: the harmonic mean of the precision and recall measures into a single score [31], [32].

$$F1\ score = \frac{2 \times Recall \times Precision}{Recall + Precision} = \frac{2TP}{2TP + FN + FP}$$

B. RADIO EXERCISE NO. 1

This appendix introduces "radio exercise no. 1," which we set as the target exercise in experiment 2. This exercise was originally designed for helping people improve their physical fitness and its current version is well known in Japan, with a history spanning more than 70 years since 1951 [47]. Because the morning assembly for this exercise is a common summer event for Japanese children, most Japanese are familiar with it and are rarely seeing it for the first time.

This exercise consists of 13 short workouts that are easily performed by ordinary people from children to the elderly, and involves exercises for the whole body including jumping and twisting/stretching/bending of the trunk. Note that the second workout and the 12th workout are exactly the same. All workouts are set to designated music with short commands on the next workout, and take approximately 3 min from start to finish [41].

Each designated workout in the official illustration [40] is shown in detail below. Note that all illustrations are cited from

TABLE 5. Confusion matrix.

Data class	Classified as positive	Classified as negative
Labeled as positive	True positive (TP)	False negative (FN)
Labeled as negative	False positive (FP)	True negative (TN)

the official illustration [40], but the written descriptions have been translated by the authors.

1. Stretching the body (Fig. 18)
2. Swinging the arms and bending/stretching the legs (Fig. 19)
3. Rotating the arms (Fig. 20)
4. Spreading the chest (Fig. 21)
5. Side bending of the body (Fig. 22)
6. Bending of the body back and forth (Fig. 23)
7. Body twisting (Fig. 24)
8. Stretching the arms up and down (Fig. 25)
9. Bending the body diagonally downwards and spreading the chest (Fig. 26)
10. Rotating the body (Fig. 27)
11. Jumping with both legs (Fig. 28)
12. Swinging the arms and bending/stretching the legs (same as the second workout)
13. Deep breaths (Fig. 29)

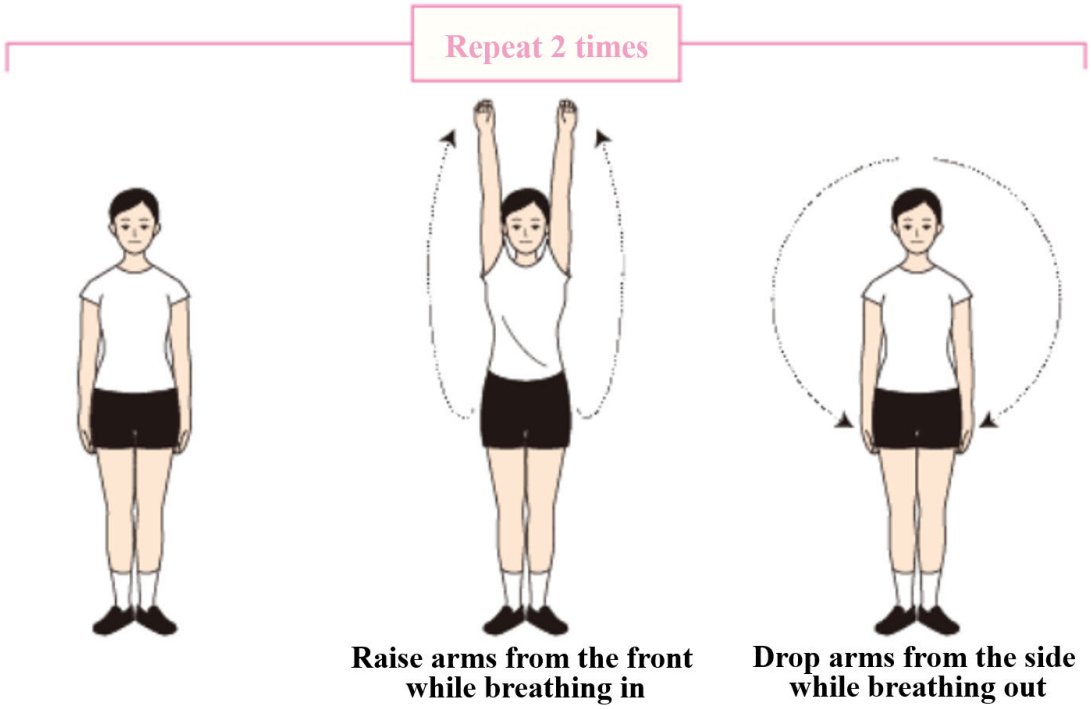

FIGURE 18. Stretching the body.

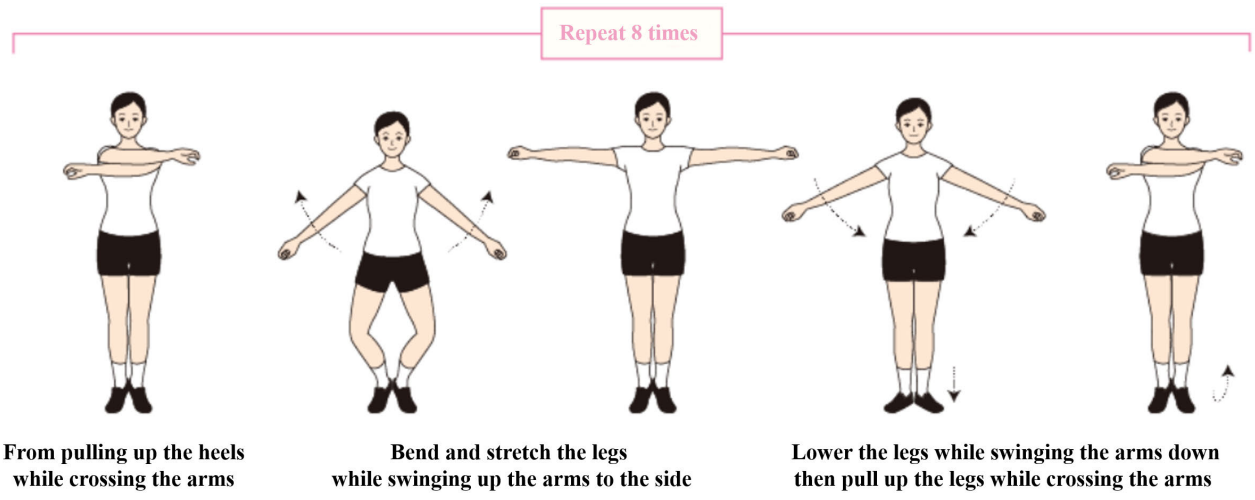

FIGURE 19. Swinging the arms and bending/stretching the legs.

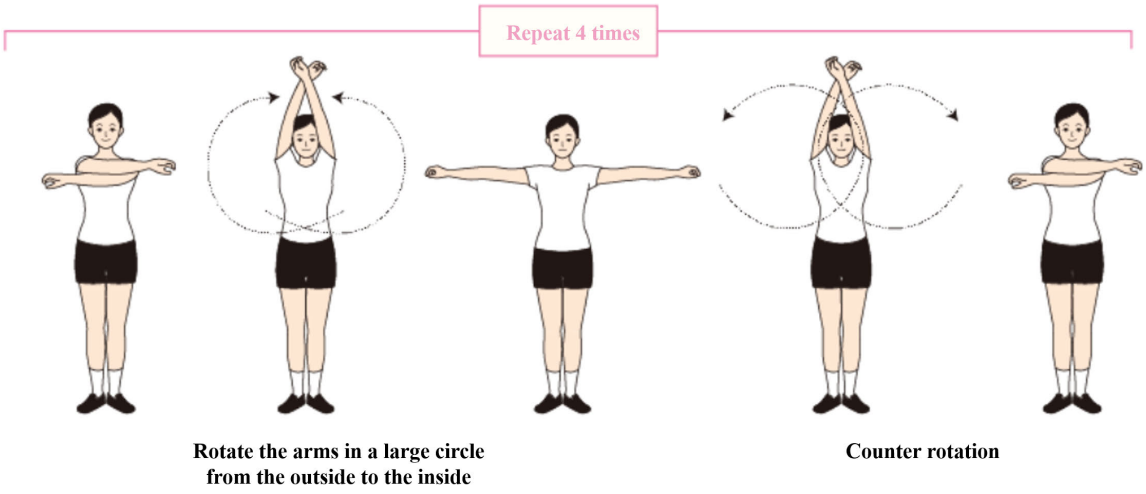

FIGURE 20. Rotating the arms.

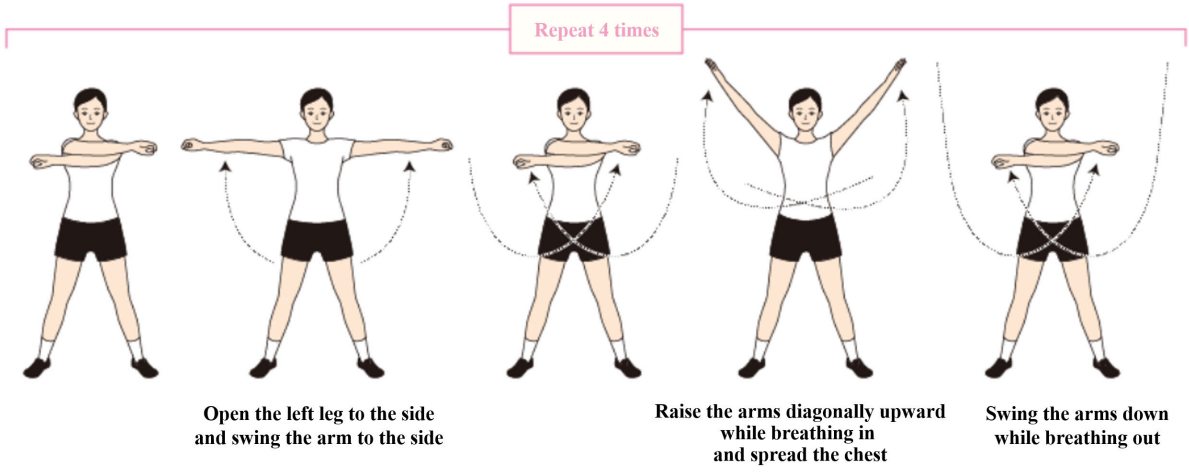

FIGURE 21. Spreading the chest.

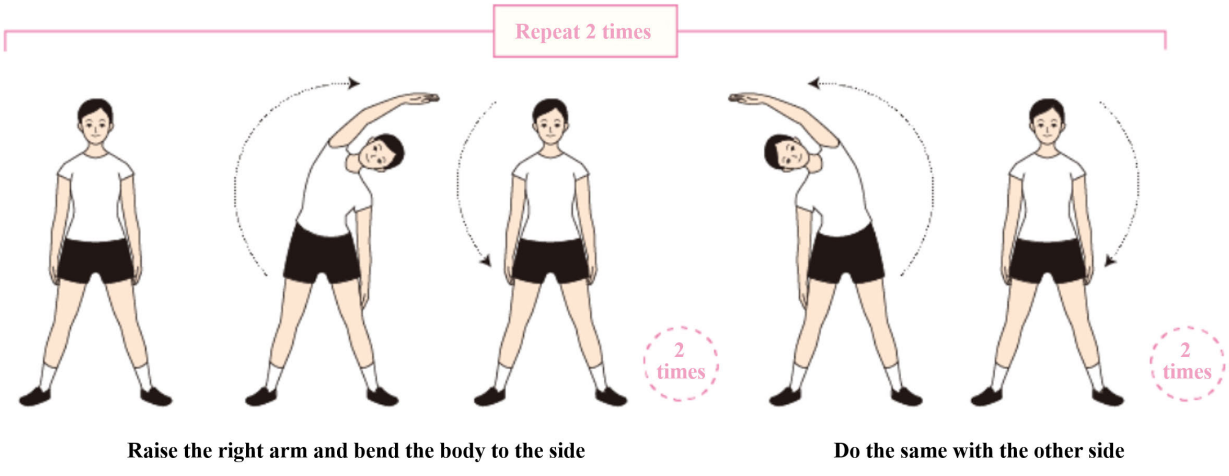

FIGURE 22. Side bending of the body.

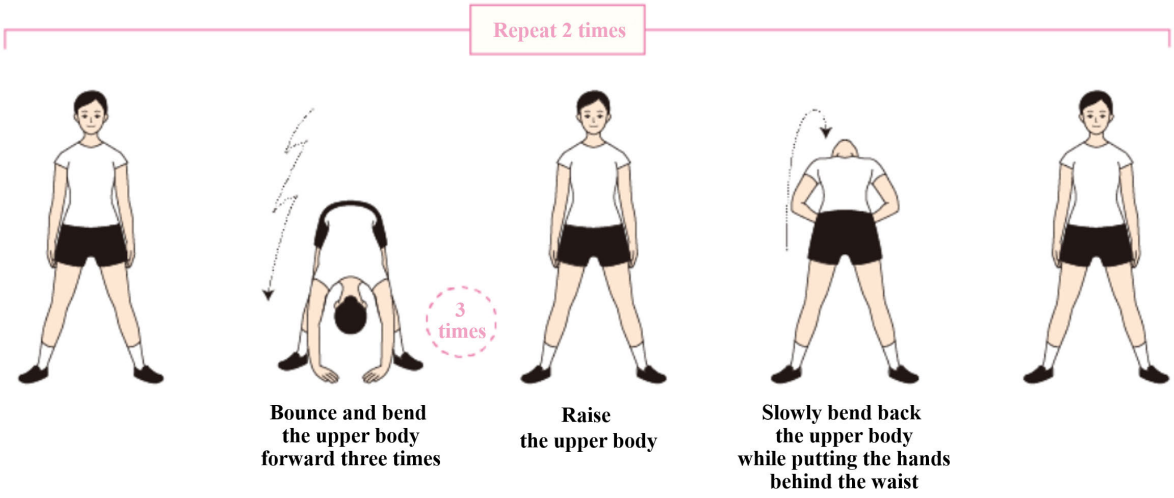

FIGURE 23. Bending of the body back and forth.

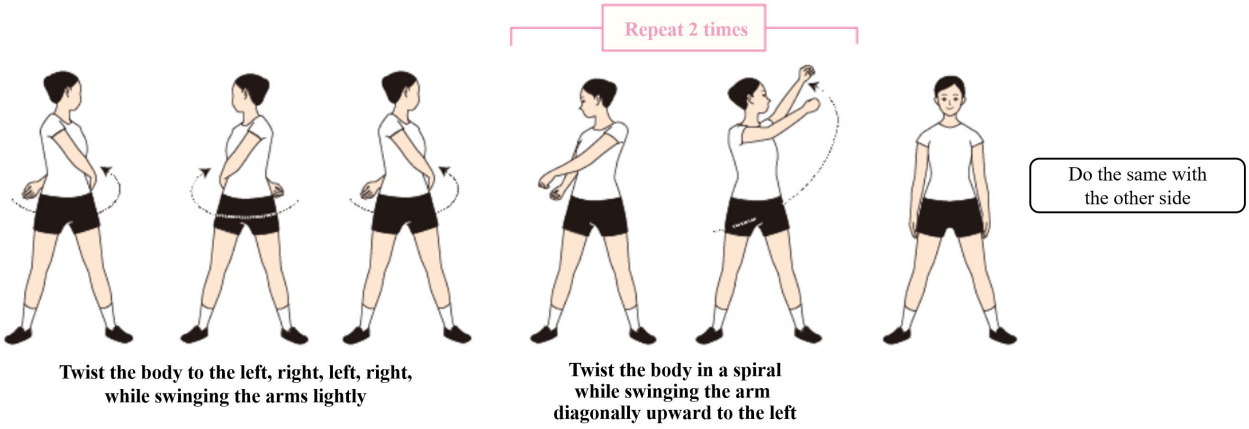

FIGURE 24. Body twisting.

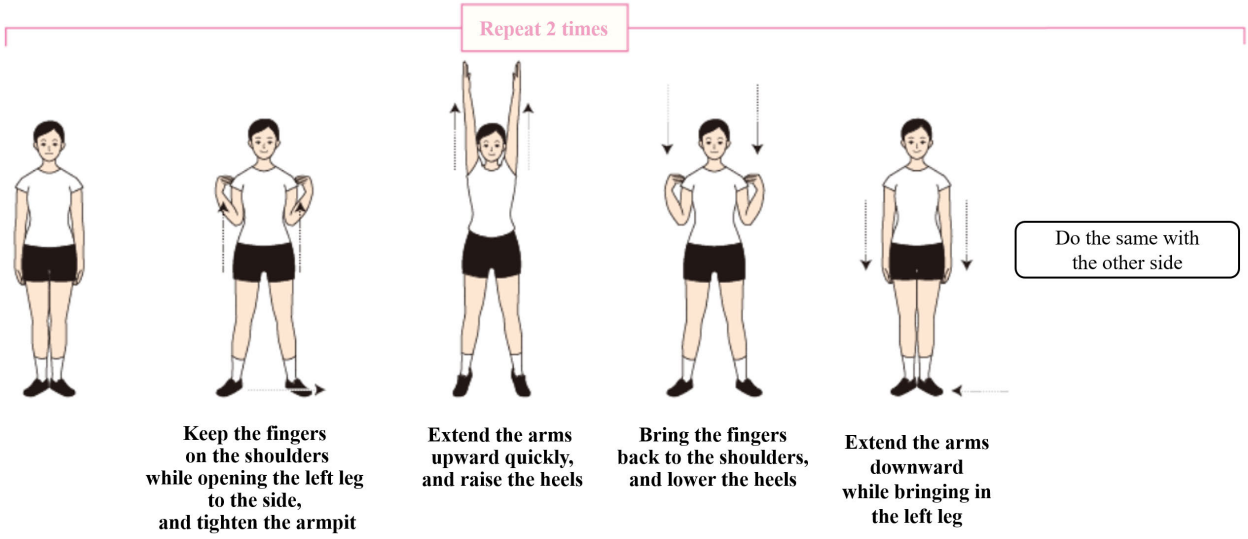

FIGURE 25. Stretching the arms up and down.

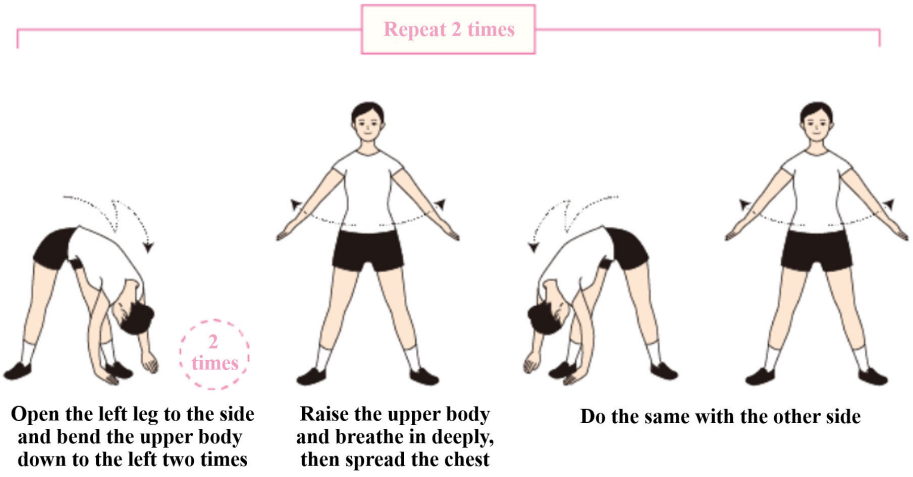

FIGURE 26. Bending the body diagonally downwards and spreading the chest.

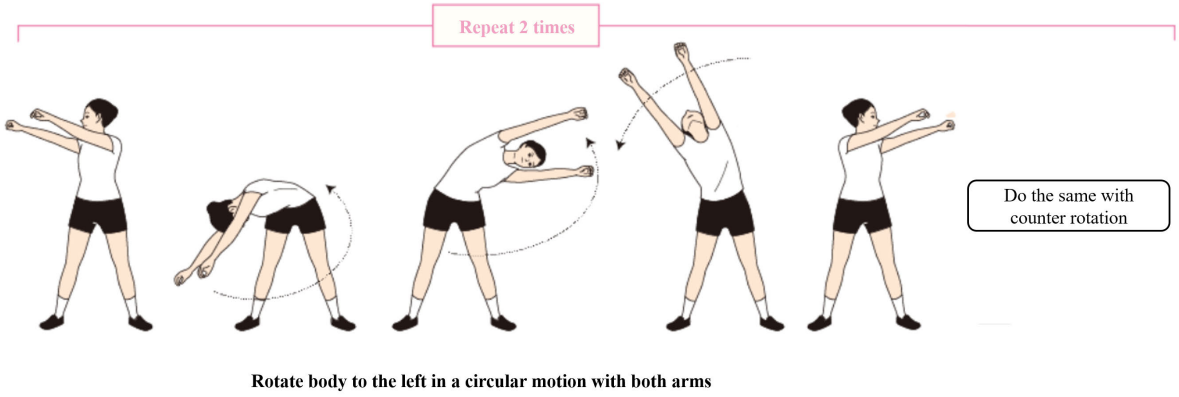

FIGURE 27. Rotating the body.

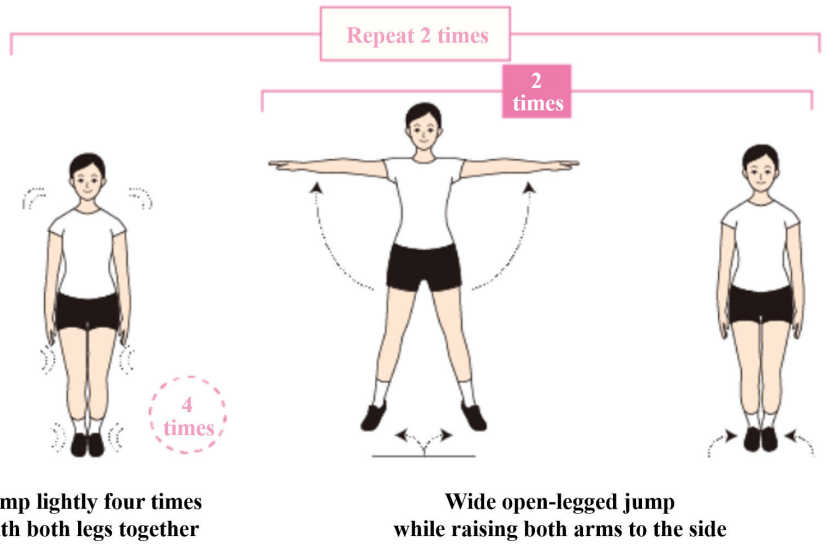

FIGURE 28. Jumping with both legs.

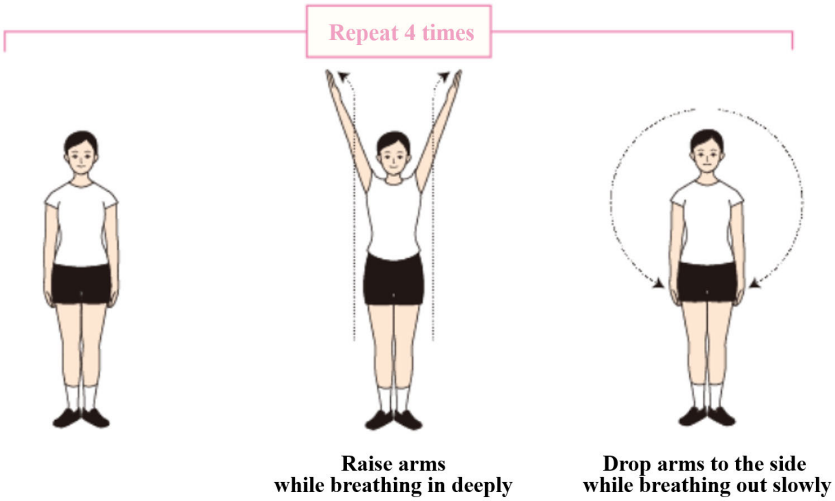

FIGURE 29. Deep breaths.

ACKNOWLEDGMENT

The authors thank Mr. Kazuhiro Yoshida at NTT DOCOMO, INC., Tokyo, Japan, for his technical support with the evaluation. They also thank Mr. Taketo Imaoka and Dr. Takuya Okuno at NTT TechnoCross Corporation, Tokyo, Japan, for their technical support on developing mobile application.

REFERENCES

[1] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, Mar. 1996, doi: 10.1093/oxfordjournals.eurheartj.a014868.

[2] T. Takeda, O. Mizuno, and T. Tanaka, "Time-dependent sleep stage transition model based on heart rate variability," in *Proc. IEEE EMBC*, Milan, Italy, Aug. 2015, pp. 2343–2346, doi: 10.1109/EMBC.2015.7318863.

[3] T. Kondo, Y. Yamato, M. Nakayama, A. Chiba, K. Sakaguchi, T. Nishiguchi, T. Masuda, and T. Yoshida, "Natural sensing with 'hitoe' functional material and initiatives towards its applications," *NTT Tech. Rev.*, vol. 15, no. 9, pp. 1–8, 2017. [Online]. Available: https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201709fa3.html

[4] K. Fujiwara, E. Abe, K. Kamata, C. Nakayama, Y. Suzuki, T. Yamakawa, T. Hiraoka, M. Kano, Y. Sumi, F. Masuda, M. Matsuo, and H. Kadotani, "Heart rate variability-based driver drowsiness detection and its validation with EEG," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1769–1778, Jun. 2019, doi: 10.1109/TBME.2018.2879346.

[5] T. Takagahara, K. Ono, N. Oda, and T. Teshigawara, "'Hitoe'—A wearable sensor developed through cross-industrial collaboration," *NTT Tech. Rev.*, vol. 12, no. 9, pp. 1–5, 2014. [Online]. Available: https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201409ra1.html

[6] Y. T. Tsukada, M. Tokita, H. Murata, Y. Hirasawa, K. Yodogawa, Y.-K. Iwasaki, K. Asai, W. Shimizu, N. Kasai, H. Nakashima, and S. Tsukada, "Validation of wearable textile electrodes for ECG monitoring," *Heart Vessels*, vol. 34, no. 7, pp. 1203–1211, Jan. 2019, doi: 10.1007/s00380-019-01347-8.

[7] N. Shiozawa, J. Lee, A. Okuno, and M. Makikawa, "Novel under wear 'smart-wear' with stretchable and flexible electrodes enables insensible monitoring electrocardiograph," in *Proc. World Eng. Conf. Conv.*, Kyoto, Japan, Nov./Dec. 2015, pp. 1–2.

[8] T. Alkhidir, A. Sluzek, and M. K. Yapici, "Simple method for adaptive filtering of motion artifacts in E-textile wearable ECG sensors," in *Proc IEEE EMBC*, Milan, Italy, Mar. 2015, pp. 3807–3810, doi: 10.1109/EMBC.2015.7319223.

[9] K. Eguchi, R. Aoki, S. Shimauchi, K. Yoshida, and T. Yamada, "R-R interval outlier processing for heart rate variability analysis using wearable ECG devices," *Adv. Biomed. Eng.*, vol. 7, pp. 28–38, Oct. 2018, doi: 10.14326/abe.7.28.

[10] T. Ogasawara, K. Ono, N. Matsuura, M. Yamaguchi, J. Watanabe, and S. Tsukada, "Development of applications for a wearable electrode embedded in inner shirt," *NTT Tech. Rev.*, no. 13, pp. 1–6, Mar. 2015. [Online]. Available: https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201501fa2.html

[11] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985, doi: 10.1109/TBME.1985.325532.

[12] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 1, pp. 42–57, Aug. 2002, doi: 10.1109/51.993193.

[13] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PLoS ONE*, vol. 9, Jan. 2014, Art. no. e84018, doi: 10.1371/journal.pone.0084018.

[14] J. P. Martinez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator: Evaluation on standard databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, Apr. 2004, doi: 10.1109/TBME.2003.821031.

[15] D. Labate, F. L. Foresta, G. Occhiuto, F. C. Morabito, A. Lay-Ekuakille, and P. Vergallo, "Empirical mode decomposition vs. Wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison," *IEEE Sensors J.*, vol. 13, no. 7, pp. 2666–2674, Jul. 2013, doi: 10.1109/JSEN.2013.2257742.

[16] P. S. Addison, "Wavelet transforms and the ECG: A review," *Physiol. Meas.*, vol. 26, no. 5, pp. R155–R199, Oct. 2005, doi: 10.1088/0967-3334/26/5/R01.

[17] B. A. Rajoub, "An efficient coding algorithm for the compression of ECG signals using the wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 4, pp. 355–362, Apr. 2002, doi: 10.1109/10.991163.

[18] L. Brechet, M.-F. Lucas, C. Doncarli, and D. Farina, "Compression of biomedical signals with mother wavelet optimization and best-basis wavelet packet selection," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2186–2192, Dec. 2007, doi: 10.1109/TBME.2007.896596.

[19] U. Yoon, I. Hwang, Y. Noh, I. Chung, and H. Yoon, "Comparison of CWT with DWT for detecting QRS complex on wearable ECG recorder," in *Proc. ICWAPR*, Quingdao, China, Jul. 2010, pp. 300–303, doi: 10.1109/ICWAPR.2010.5576361.

[20] C. Zeng, H. Lin, Q. Jiang, and M. Xu, "QRS complex detection using combination of mexican-hat wavelet and complex Morlet wavelet," *J. Comput.*, vol. 8, no. 11, pp. 2951–2958, Nov. 2013.

[21] A. M. Hamad Alhussainy, "QRS complex detection and RR interval computation based on discrete wavelet transform," *Int. J. Smart Sens. Intell. Syst.*, vol. 13, no. 1, pp. 1–11, 2020, doi: 10.21307/ijssis-2020-010.

[22] R. Aoki, K. Eguchi, S. Shimauchi, K. Yoshida, and T. Yamada, "Consideration of calculation process assuming heart rate variability analysis using wearable ECG devices," in *Proc EMBC*, Honolulu, HI, USA, 2018, pp. 5693–5696, doi: 10.1109/EMBC.2018.8513449.

[23] M. Jia, F. Li, J. Wu, Z. Chen, and Y. Pu, "Robust QRS detection using high-resolution wavelet packet decomposition and time-attention convolutional neural network," *IEEE Access*, vol. 8, pp. 16979–16988, 2020, doi: 10.1109/ACCESS.2020.2967775.
[24] M. U. Gul, K. Kadir, H. K. Azman, and S. Iqbal, "Detection of R-peaks using single-scale wavelet transform," in *Proc MACS*, Karachi, Pakistan, 2019, pp. 1–5, doi: 10.1109/MACS48846.2019.9024782.
[25] S. Shimauchi, K. Eguchi, T. Takeda, and R. Aoki, "An analysis method for wearable electrocardiogram measurement based on non-orthogonal complex wavelet expansion," in *Proc IEEE EMBC*, Seogwipo, South Korea, Dec. 2017, pp. 3973–3976, doi: 10.1109/EMBC.2017.8037726.
[26] L. Senhadji, L. Thoraval, and G. Carrault, "Continuous wavelet transform: ECG recognition based on phase and modulus representations and hidden Markov models," in *WAVELETS in Medicine and Biology*, A. Aldroubi and M. A. Unser Eds. Boca Raton FL, USA: CRC Press, 1996, pp. 439–463.
[27] P. S. Addison, J. N. Watson, and T. Feng, "Low-oscillation complex wavelets," *J. Sound Vib.*, vol. 254, no. 4, pp. 733–762, Jul. 2002, doi: 10.1006/jsvi.2001.4119.
[28] J. Okude, *Korenara Wakaru! Kantan Piont Shinden-Zu (in Japanese) [Easy-to-Understand Tutorial on Electrocardiogram]*, 2nd ed. Tokyo, Japan: Igaku-Shoin, 2011.
[29] G. Clifford, F. Azuaje, and P. McSharry, *Advanced Methods And Tools for ECG Data Analysis*, 1st ed. Norwood, MA, USA: Artech House, 2006.
[30] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.
[31] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.
[32] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*. Newton, MA, USA: O'Reilly Media, 2017.
[33] H. Kim, R. F. Yazicioglu, P. Merken, C. Van Hoof, and H.-J. Yoo, "ECG signal compression and classification algorithm with quad level vector for ECG holter system," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 1, pp. 93–100, Jan. 2010, doi: 10.1109/TITB.2009.2031638.
[34] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, "Kubios HRV–heart rate variability analysis software," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 210–220, Jan. 2014, doi: 10.1016/j.cmpb.2013.07.024.
[35] G. B. Moody, W. E. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," *Comput. Cardiol.*, vol. 11, pp. 381–384, 1984. [Online]. Available: http://ecg.mit.edu/george/publications/nst-cinc-1984.pdf
[36] Physionet. *The MIT-BIH Noise Stress Test Database*. Accessed: Dec. 4, 2020. [Online]. Available: https://archive.physionet.org/physiobank/database/nstdb/
[37] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001, doi: 10.1109/51.932724.
[38] Physionet. *MIT-BIH Arrhythmia Database*. Accessed: Dec. 4, 2020. [Online]. Available: https://archive.physionet.org/physiobank/database/mitdb/
[39] Physionet. *PhysioBank Annotations*. Accessed: Dec. 4, 2020. [Online]. Available: https://archive.physionet.org/physiobank/annotations.shtml
[40] Japan Post Insurance. *Radio Exercise No. 1 (In Japanese)*. Accessed: Dec. 4, 2020. [Online]. Available: https://www.jp-life.japanpost.jp/radio/instruction/radio_first.html
[41] Japan Post Insurance. *Radio Exercise No. 1 (In Japanese)*. Accessed: Dec. 4, 2020. [Online]. Available: https://www.youtube.com/watch?v=_YZZfaMGEOU
[42] S. Tsukada, H. Nakashima, and K. Torimitsu, "Conductive polymer combined silk fiber bundle for bioelectrical signal recording," *PLoS ONE*, vol. 7, no. 4, Apr. 2012, Art. no. e33689, doi: 10.1371/journal.pone.0033689.
[43] H. Blackburn, H. L. Taylor, C. L. Vasquez, and T. C. Puchner, "The electrocardiogram during exercise: Findings in bipolar chest leads of 1,449 middle-aged men, at moderate work levels," *Circulation*, vol. 34, no. 6, pp. 1034–1043, Dec. 1966, doi: 10.1161/01.cir.34.6.1034.
[44] K. Eguchi, R. Aoki, K. Yoshida, and T. Yamada, "Reliability evaluation of R-R interval measurement status for time domain heart rate variability analysis with wearable ECG devices," in *Proc. IEE EMBC*, Seogwipo, South Korea, Oct. 2017, pp. 1307–1311, doi: 10.1109/EMBC.2017.8037072.
[45] K. Eguchi, R. Aoki, K. Yoshida, and T. Yamada, "R-R interval outlier exclusion method based on statistical ECG values targeting HRV analysis using wearable ECG devices," in *Proc. IEEE EMBC*, Honolulu, HI, USA, Dec. 2018, pp. 5689–5692, doi: 10.1109/EMBC.2018.8513452.
[46] K. Kamata, K. Kinoshita, and M. Kano, "Missing RRI interpolation algorithm based on locally weighted partial least squares for precise heart rate variability analysis," *Sensors*, vol. 18, no. 11, p. 3870, Nov. 2018, doi: 10.3390/s18113870.
[47] Japan Post Insurance. *Radio Exercise/Exercises for Everyone (in Japanese)*. Accessed: Dec. 4, 2020. [Online]. Available: https://www.jp-life.japanpost.jp/radio/index.html

SUEHIRO SHIMAUCHI (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1991, 1993, and 2007, respectively.

He was with Nippon Telegraph and Telephone (NTT) Corporation, Tokyo, in 1993, where he engaged in research on acoustic and biomedical signal processing. He joined the Kanazawa Institute of Technology, Ishikawa, Japan, in 2018, where he is currently a Professor.

Dr. Shimauchi is a member of the Institute of Electronics, Information, and Communication Engineers of Japan (IEICE), and the Acoustic Society of Japan (ASJ).

KANA EGUCHI (Member, IEEE) received the B.E. degree from the Kyoto Institute of Technology, Kyoto, Japan, in 2010, and the M.S. and Ph.D. degrees in informatics from Kyoto University, Kyoto, in 2012 and 2020, respectively.

She joined NTT Corporation, Tokyo, Japan, in 2012. She has been engaged in research on network middleware, medical engineering, and medical informatics at NTT Service Evolution Laboratories, Kanagawa, Japan. Her current research interests include biosignal processing, wearable/ubiquitous computing, medical engineering, and medical informatics.

Dr. Eguchi is a member of the IEICE, the IEEE Engineering in Medicine and Biology Society (EMBS), the Japanese Society for Medical and Biological Engineering (JSMBE), the International Society of Electrophysiology and Kinesiology (ISEK), and the Japanese Society of Sleep Research (JSSR). Her awards include Telecom System Technology Award for Student (Honorable Mention) from the Telecommunications Advancement Foundation, in 2019.

RYOSUKE AOKI received the B.E., M.S., and Ph.D. degrees in information sciences from Tohoku University, Miyagi, Japan, in 2005, 2007, and 2014, respectively.

He joined NTT Corporation, Tokyo, Japan, in 2007. He is currently working at NTT Service Evolutions Laboratories, Kanagawa, Japan. His current research interests include human–computer interaction, interaction design, research through design, medical engineering, and wearable/ubiquitous computing.

Dr. Aoki is a member of the Information Processing Society of Japan (IPSJ).

MASAHIRO FUKUI (Member, IEEE) received the B.E. degree in information science from Ritsumeikan University, Shiga, Japan, in 2002, the M.E. degree in information science from the Nara Institute of Science and Technology, Nara, Japan, in 2004, and the Ph.D. degree in information science from Ritsumeikan University, in 2018.

Since joining NTT Corporation, in 2004. He has been engaged in research on acoustic echo cancellers and speech coding. He is currently a Senior Research Engineer at NTT Media Intelligence Laboratories.

Dr. Fukui is a member of the ASJ. His awards include the Best Paper Award of the IEEE International Conference on Consumer Electronics (ICCE) Conference and the Technical Development Award from the ASJ, in 2014.

NOBORU HARADA (Senior Member, IEEE) received the B.S. and M.S. degrees in computer science from the Kyushu Institute of Technology, Fukuoka, Japan, in 1995 and 1997, respectively, and the Ph.D. degree in computer science from the University of Tsukuba, Ibaraki, Japan, in 2017.

Since joining NTT Corporation, Tokyo, Japan, in 1997, he has been involved with research on speech and audio signal processing, such as high efficiency coding and lossless compression. His current research interests include acoustic signal processing and machine learning for acoustic event detection, including anomaly detection in sound.

Dr. Harada is a member of ASJ, IEICE, and IPSJ. He was a recipient of the Technical Development Award from the ASJ, in 2016, the Industrial Standardization Encouragement Award from the Ministry of Economy Trade and Industry of Japan, in 2011, and Telecom System Technology Award (Encouragement Award) from the Telecommunications Advancement Foundation, in 2007 and 2020.