



Abnormal beat detection from unreconstructed compressed signals based on linear approximation in ECG signals suitable for embedded IoT devices

Seungmin Lee¹ · Daejin Park¹

Received: 8 January 2021 / Accepted: 25 October 2021 / Published online: 1 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In the study of electrocardiogram signal monitoring systems, signal compression techniques for effective signal transmission and abnormal beat detection for arrhythmia diagnosis are paramount areas. The general abnormal beat detection has a problem of unsuitability for low-power and low-capacity embedded devices because it requires reconstructing a compressed signal to generate an auxiliary signal for fiducial point (FP) detection. In this study, we propose a method to compress signals into a small number of vertices, including the FP, using an optimized dynamic programming-based linear approximation. Then, we detect the FP from the vertices and classify the abnormal beat. The proposed method minimizes the amount of memory usage and computation by detecting the FP using the vertex's feature value without reconstructing the compressed signal. The signal compression performance of the proposed method showed an average compression ratio of 7.05:1 and a root-mean-square difference of 0.78% for 48 records of the MIT-BIH arrhythmia database. In addition, the premature ventricular contraction abnormal beat detection performance using only QR interval feature showed 84.07% sensitivity and 93.70% accuracy; when R-peak's amplitude and RR interval features were added, the sensitivity and accuracy increased to 96.65% and 93.76%, respectively. Therefore, we confirmed that the proposed method could effectively compress electrocardiogram signals based on linear approximation and detect abnormal beats without signal reconstruction.

Keywords ECG · Signal compression · Linear approximation · Signal reconstruction · QRS interval · Premature ventricular contraction

1 Introduction

Biosignal monitoring acquires and transmits signals, detects abnormal signals, and checks the user's health in real time (Meng et al. 2019; Lee et al. 2019a). The electrocardiogram (ECG) signal is a representative biosignal used for the early diagnosis of heart disease, and biosignal monitoring studies on the ECG signal are being actively conducted as the mortality rate due to heart disease increases (Arshad et al. 2016). In addition, it is used for smart home elderly care, security, human activity, and so on (Rhim et al. 2018; Teraoka 2012).

The abnormal beat used for an arrhythmia diagnosis of the ECG signal requires a long-time measurement because its frequency of occurrence is low. In addition, feature values, such as amplitudes, intervals, and segments, are acquired based on the detection of fiducial points (FPs), such as the onset (start of waveform) and offset (end of waveform) of each waveform, and the abnormal beat is classified based on these feature values (Lee and Park 2021b; Lee et al. 2017; Yazdani and Vesin 2016; Lin et al. 2014; Martinez et al. 2010; Illanes-Manriquez and Zhang 2008; Madeiro et al. 2012; Laguna et al. 1994). Therefore, a high sampling frequency is required for accurate FP detection.

To store and transmit a signal acquired for a long time at a high sampling frequency, an effective signal compression technique is required. Signal compression is classified into frequency- and time-domain compression; wavelet-based signal compression (Bera et al. 2020) is a representative frequency-domain compression method. The linear approximation method is a time-domain signal compression

✉ Daejin Park
boltanut@knu.ac.kr

¹ School of Electronic and Electrical Engineering, Kyungpook National University, 80 Daehak-Ro, Buk-Gu, Daegu 41566, Korea

technique in which a small number of vertices represent a signal (Lee and Park 2021a; Lee et al. 2019b). In particular, linear approximation includes FPs as vertices and enables effective detection by emphasizing the feature values of the FPs. In addition, as FPs can be detected by analyzing the vertices' feature values obtained through linear approximation, it has the advantage of not having to detect FPs after reconstructing the compressed signal.

In this article, a method that enables the compression of ECG signals effectively based on the characteristics of linear approximation and detection of abnormal beats without signal reconstruction from an approximated signal is proposed. ECG signals are periodically sampled, so they consist of a large amount of amplitude information. The amplitude information is the same as that of the input signal, but additional time information is required because the vertices of an approximated signal are aperiodic. The time information on the vertices of an approximated signal requires a large amount of data because the ECG signal is measured for a long time, which significantly reduces the compression rate (Lee and Park 2020). In this article, we improve signal compression rate by storing the time difference from the previous vertex instead of the time information of the vertex based on the constraint of the time difference between the vertices used in the linear approximation.

Further, we acquire the width of the QRS complex using the FPs of the QRS complex, which are obtained from the approximated vertices. By confirming that the proposed method enables the detection of abnormal beats, including premature ventricular contractions (PVCs), which increase the width of the QRS complex, we confirmed that the proposed method enables FP detection and abnormal beat detection without signal reconstruction from the approximated signal. Figure 1 shows the flowchart of the proposed method.

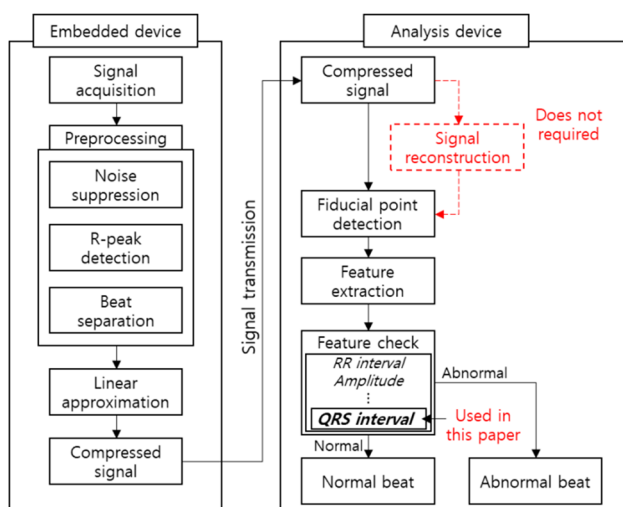


Fig. 1 Flowchart of the proposed method

The rest of this article is organized as follows. Section 2 introduces the composition of ECG signals and characteristics of abnormal beats. Section 3 introduces the process of linear approximation and FP detection in detail. Section 4 introduces the proposed signal compression method, followed by the abnormal beat detection method. In Sect. 5, we assess the performance of the proposed method experimentally. Finally, we conclude this study in Sect. 6.

2 ECG signal

2.1 Composition of ECG signal

During depolarization and repolarization of the atrium and ventricle, the P-wave, QRS complex, and T-wave are measured in the ECG signal (Huszar 2007). The feature values of each waveform are obtained using the intervals, segments, and amplitude difference between FPs.

Figure 2 shows the composition of the FPs and feature values of an ECG signal.

Among FPs, the R-peak has the highest amplitude value. Therefore, R-peak detection is relatively easy and can be accomplished accurately through postprocessing correction (Pan and Tompkins 1985; James 2015). Based on these characteristics, the R-peak becomes a criterion for classifying beats and is used for detecting other FPs.

Each beat is classified through the R-peak; two methods are generally used for separating the beats. The first method divides the beats into the region of 275 ms before and 375 ms after the R-peak (Li et al. 2010). This R-peak centered region includes the P-wave, QRS complex, and T-wave, which are mainly used for ECG signal analyses. However, some data samples are missing. The second method divides the beats using the RR interval, thereby minimizing the approximation error by approximating all data. Figure 3 shows the two methods of beat separation.

In this article, the beats are separated according to the RR interval (Fig. 3b).

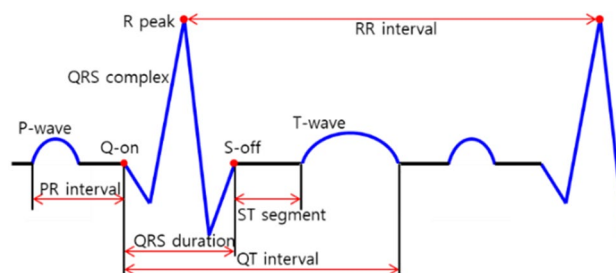


Fig. 2 Composition of FPs and feature values of ECG signal

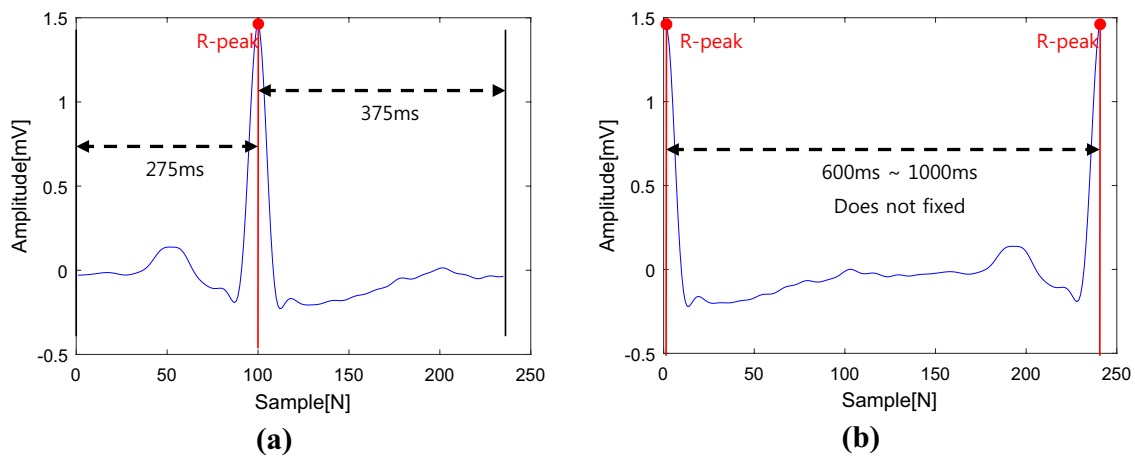


Fig. 3 Beat separation methods: **a** centering on the R-peak and **b** RR interval

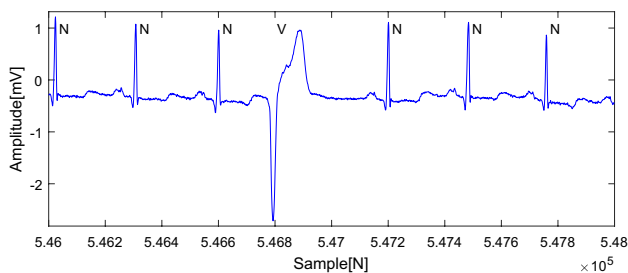


Fig. 4 Example of PVC

2.2 Characteristics of abnormal beat

Abnormal beats due to arrhythmia appear as a change in the shape or period of a waveform in an ECG signal. Figure 4 depicts an example of a typical abnormal beat: PVC.

As shown in Fig. 4, PVC (V) has a short RR interval and deformed QRS complex compared with a normal beat (N). The variation of the RR interval can be determined through the R-peak detection. To detect a deformed QRS complex, various features, such as the amplitude and kurtosis of the R-peak and the width of the QRS complex, are used, or the abnormal beats are classified by learning the normal beat's shape (Ou et al. 2020). In addition, a deep learning method using convolutional neural network architecture has been used for abnormal detection (Ozdemir et al. 2020).

In this study, we employed the width of the QRS complex to use the information on the FP for the onset and offset of the QRS complex acquired via linear approximation.

2.3 FP detection

Abnormal beat classification is divided into shape- and feature-based classifications. Feature-based classification requires each waveform's FPs (Fig. 2). As mentioned above,

among the FPs, the R-peak has the largest amplitude value in the beat, so it is easy to detect, and the detection is reliable. A representative R-peak detection method is Pan's method. The R-peak detection results are used to classify beats, and the amplitude and RR interval of the R-peak are used for abnormal beat detection.

Based on the R-peak detection results, various FP detection methods, such as the onset and offset of the QRS complex and P-wave are being studied. However, unlike the R-peak, the onset and offset of the QRS complex and P-wave are non-local extremum points, so they are ambiguous and difficult to detect. To detect FPs, an auxiliary signal is acquired using a method, such as differentiation (Laguna et al. 1994) or Hilbert transformation (Manriquez and Zhang 2007). Thus, it is necessary to reconstruct the compressed signal because the existing FP detection methods require an auxiliary signal. In addition, there is a problem of having a different FP detection result from that of the input signal due to the reconstruction error.

To mitigate the abovementioned issues, we propose an efficient signal compression method using linear approximation. In addition, we confirm that abnormal beats can be detected by detecting FPs and acquiring the feature values without signal reconstruction based on the characteristics of the FP detection using only the approximated vertices.

An embedded device acquires an ECG signal, compresses the preprocessed signal through linear approximation, and transmits the signal. The preprocessing consists of noise suppression to detect a stable beat (Mohd Apandi et al. 2020), R-peak detection to separate the beats, and beat separation according to the RR interval (Fig. 2b). The analysis device detects FPs and acquires feature values without signal reconstruction from the approximated signal as well as analyzes feature values to detect abnormal beats. In this article, we confirm that abnormal beats can be detected from an approximated signal using the width of the QRS complex

obtained from the detection of the onset and offset of the QRS complex.

3 Linear approximation

An approximation is a representative signal compression method in the time domain; linear approximation and B-spline are typical methods. Approximations effectively express a signal using a small number of vertices. B-spline is widely used because it guarantees signal continuity and has a low reconstruction error; however, the amount of computation and storage usage significantly increase, making it difficult to implement in embedded devices. In addition, B-spline requires FP detection after signal reconstruction because the location of knots does not coincide with FPs. By contrast, the linear approximation method is simple to implement in an embedded device but also represents the FP that is the boundary between the baseline and the waveform as a vertex.

Figure 5 shows the concept of the linear approximation method.

As shown in Fig. 5a, the detection of FPs, such as the onset or offset, is difficult because the FPs do not have a

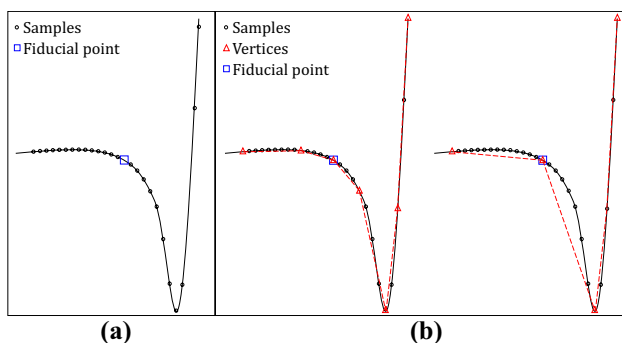
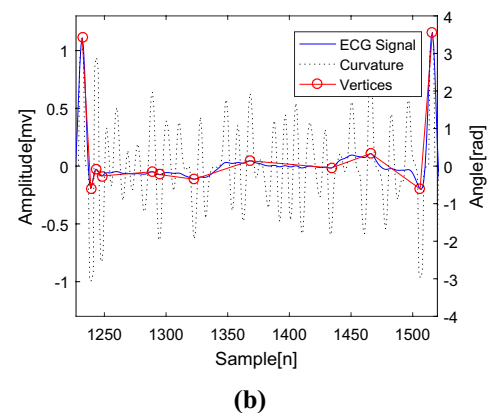
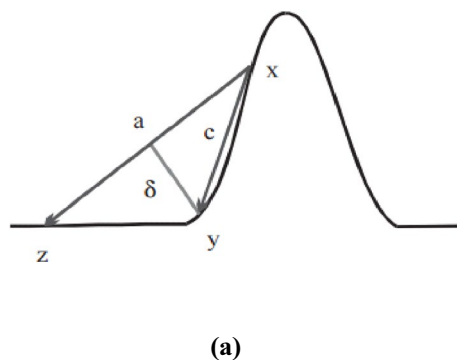


Fig. 5 Illustration of linear approximation: **a** existing method and **b** linear approximation

Fig. 6 Curvature-based linear approximation: **a** concept of curvature and **b** example of an approximated signal



maximum or minimum amplitude. In addition, it is ambiguous to determine an FP because the samples around FPs have similar feature values.

However, linear approximation has the advantage of representing the FP corresponding to the boundary between the waveform region with large amplitude change and the baseline region with small amplitude change as a vertex (Fig. 5b) because it represents a signal with a small number of vertices. In addition, it is possible to emphasize the feature values of vertices by representing them as a few vertices.

The input signal is divided into RR intervals based on R-peak detection, and an independent linear approximation is performed for each beat. The linear approximation proposed by Lee et al. (2018) proceeds in three stages—initial vertex selection, additional vertex selection, and error optimization. In addition, Lee et al. (2019c) and Lee and Park (2020) improved dynamic programming used in the optimization step according to the ECG signal's characteristics.

3.1 Initial vertex selection

First, a point with a large curvature is obtained as an initial vertex using a curvature-based linear approximation (Mokhtarian and Suomela 1998). The curvature is the amount by which a curve deviates from being a straight line, and it is calculated as (1) using the included angle between three points:

$$\delta(Y;X,Z) = \frac{\vec{a} \times \vec{c}}{|\vec{a}|} = \frac{(x_1 - z_1)(x_2 - y_2) - (x_2 - z_2)(x_1 - y_1)}{\sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2}}, x_1 > y_1 > z_1. \quad (1)$$

Figure 6 shows the concept of the curvature calculation process and an example of the approximated signal.

3.2 Additional vertex selection

Curvature-based linear approximation well-expresses an FP as an initial vertex because it is a point with a large curvature. However, if the waveform's amplitude change is low, it might not be acquired as the initial vertex, as shown in QRS-onset and P-offset in Fig. 6b. To correct this, additional vertices are obtained using a sequential linear approximation for the inside of each initial vertex (O'Connell 1997).

Figure 7 shows the concept of the sequential linear approximation process and an example of the approximated signal of Fig. 6b.

As shown in Fig. 7a, a vertex is added when the error exceeds the threshold D_{max} , and the signal was effectively approximated (Fig. 7b).

3.3 Error optimization

Global error optimization calculates the costs for all cases and detects the minimum case as an optimization result. Therefore, the problem of optimizing the position of N vertices for L samples generally has a complexity of $O(L^N)$. Dynamic programming (Bellman and Dreyfus 2015) is a global optimization technique in which the optimal path between two points is optimized based on Bellman's optimal principle as the globally optimal path between any two points on the global path. The recursive approach simplifies and optimizes the problem, especially using memoization to remember the computational results, which eliminates redundant operations. Dynamic programming enables high-speed global optimization but requires additional memory for memoization, which consists of cost and base matrices. In this case, the size of the cost and base matrices required for memoization is $O(L^2N)$.

Figure 8 shows the cost and base matrices used for memoization of dynamic programming.

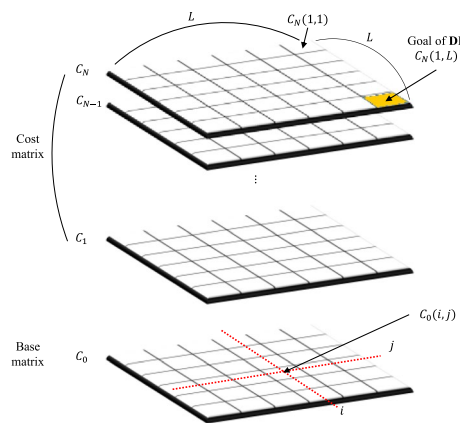


Fig. 8 Cost and base matrices for memoization of dynamic programming

For a signal with length L , the optimization of the partial signal from i to j , including the k vertices, is recursively calculated as follows:

$$C_k(i, j) = \min_{v_k \in [1, \dots, L]} (C_{k-1}(i, v_k) + C_0(v_k, j)), \quad (2)$$

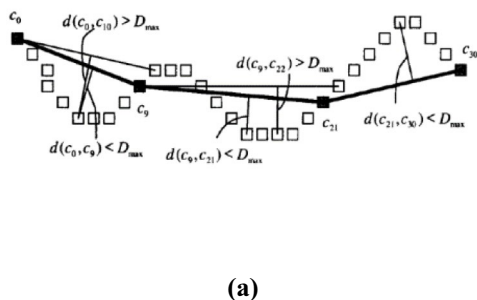
where v_k denotes the position of the k th vertex; when $k = 0$, $C_0(i, j)$ is calculated as an error between the input signal and the line connecting the i th and j th samples.

Figure 9 shows the results of optimization using the dynamic programming shown in Fig. 7b).

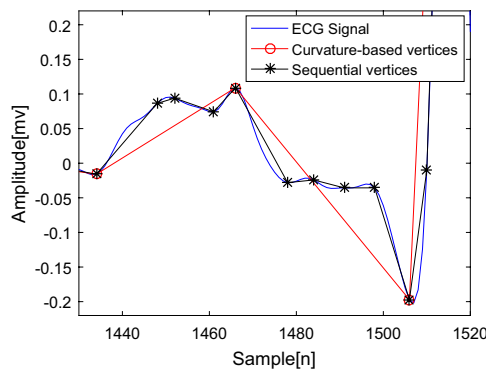
As shown in Fig. 9, the signal is well represented with a small number of vertices, and dynamic programming reduces the reconstruction error.

3.4 Optimization of dynamic programming

Lee et al. (2019c) optimized the calculation of the cost and base matrices based on ECG signal characteristics, thereby



(a)



(b)

Fig. 7 Sequential linear approximation: a concept of method, and b approximation result of Fig. 6b

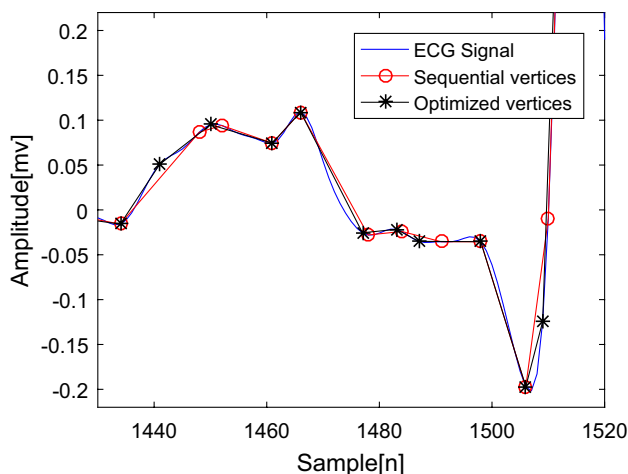


Fig. 9 Optimization results of Fig. 7(b)

improving the performance and enabling real-time operation even in embedded devices.

3.5 Improvement based on characteristics of ECG signals

The first and last vertices are fixed as initial vertices corresponding to both ends of the input signal. Therefore, in $C_k(i, j)$ of (2), i always becomes 1, and the existing cost matrix $C_k(i, j)$ can be expressed as $C(k, j)$, as follows:

$$C(k, j) = \min_{1 < v_k < j} (C(k - 1, v_k) + C_0(v_k, j)) \tag{3}$$

By expressing $C_k(1, j)$ as $C(k, j)$, the cost matrix of the size of $O(L^2N)$ is reduced to the size of $O(NL)$. Thus, $C(N, L)$ is an optimized error value when N additional vertices exist between the first and L th last samples, which represents the result of dynamic programming. In addition, the existing dynamic programming method optimizes by a recursive, top-down calculation method. However, improving the cost matrix fixes the area required for the calculation. Thus, the optimization can be done through a bottom-up operation without a recursive approach.

3.6 Adding limit of time difference between vertices

The amount of information is minimized by representing the vertex information as a time difference from the previous vertex. The maximum interval between the vertices is limited by the number of bits ($N_{Bit} = 2^5$ in this study) to indicate the time difference between the vertices. The cost matrix

$C(k, j)$ is calculated only when the interval between the k th and $(k + 1)$ th vertices does not exceed N_{Bit} , as follows:

$$C(k, j) = \min_{v_{k+1} - N_{Bit} \leq v_k < v_{k+1}} (C(k - 1, v_k) + C_0(v_k, j)). \tag{4}$$

The computation for the base matrix is reduced because the time difference between the two vertices is limited by N_{Bit} .

3.7 Row-wise operation

The base matrix is sequentially called column by column when each component of the cost matrix is calculated in a column-wise operation, i.e., the j th column of the base matrix is used only in the calculation of the cost matrix's j th column. Thus, the base matrix $C_0(v_k, j)$ in (4) can be expressed as follows:

$$C(k, j) = \min_{v_{k+1} - N_{Bit} \leq v_k < v_{k+1}} (C(k - 1, v_k) + C_0^j(v_k - (v_{k+1} - N_{Bit}) + 1)). \tag{5}$$

As shown in (5), the memory usage of the base matrix C_0^j can be overwritten by C_0^{j+1} . Accordingly, the size of the base matrix can be minimized from $O(L^2)$ to $O(N_{Bit})$ in a column unit vector.

3.8 Auxiliary signal

Without signal reconstruction, FPs are detected using the feature values of the approximated vertices. In this study, we detect the onset and offset of the QRS complex. The QRS interval is acquired from FPs, and abnormal beats are detected using this feature.

Various shapes of the QRS complex, such as downward waveform and the presence or absence of the Q- and S-peaks, make it difficult to detect FPs. To reduce the difficulty of extracting features from the QRS complex's ambiguous shape, robust FP detection methods using various auxiliary signals, such as differentiating the signal, average filtering, and Hilbert transformation, have been proposed (Illanes-Manriquez and Zhang 2008; Pan and Tompkins 1985). Lee et al. (2018) used the cumulative signal of an approximated signal to preserve the morphological features of the vertices that represent FPs.

First, we obtain the amplitude difference (V^D), as follows:

$$V^D = \{v_1^D, \dots, v_{N_V}^D\}, v_i^D = (v_{x_i}^D, v_{y_i}^D) \tag{6}$$

$$v_{x_i}^D = v_{x_i}, v_{y_i}^D = v_{y_i} - v_{y_{i-1}}, v_{y_1}^D = 0.$$

By accumulating the absolute value of the amplitude difference in (6), we generate the cumulative signal, as follows:

$$V^{D'} = \{v_1^{D'}, \dots, v_{N_V}^{D'}\}, v_i^{D'} = (v_{x_i}^{D'}, v_{y_i}^{D'}) \tag{7}$$

$$v_{x_i}^{D'} = v_{x_i}^D, v_{y_i}^{D'} = \sum_{k=1}^i |v_{y_k}^D|.$$

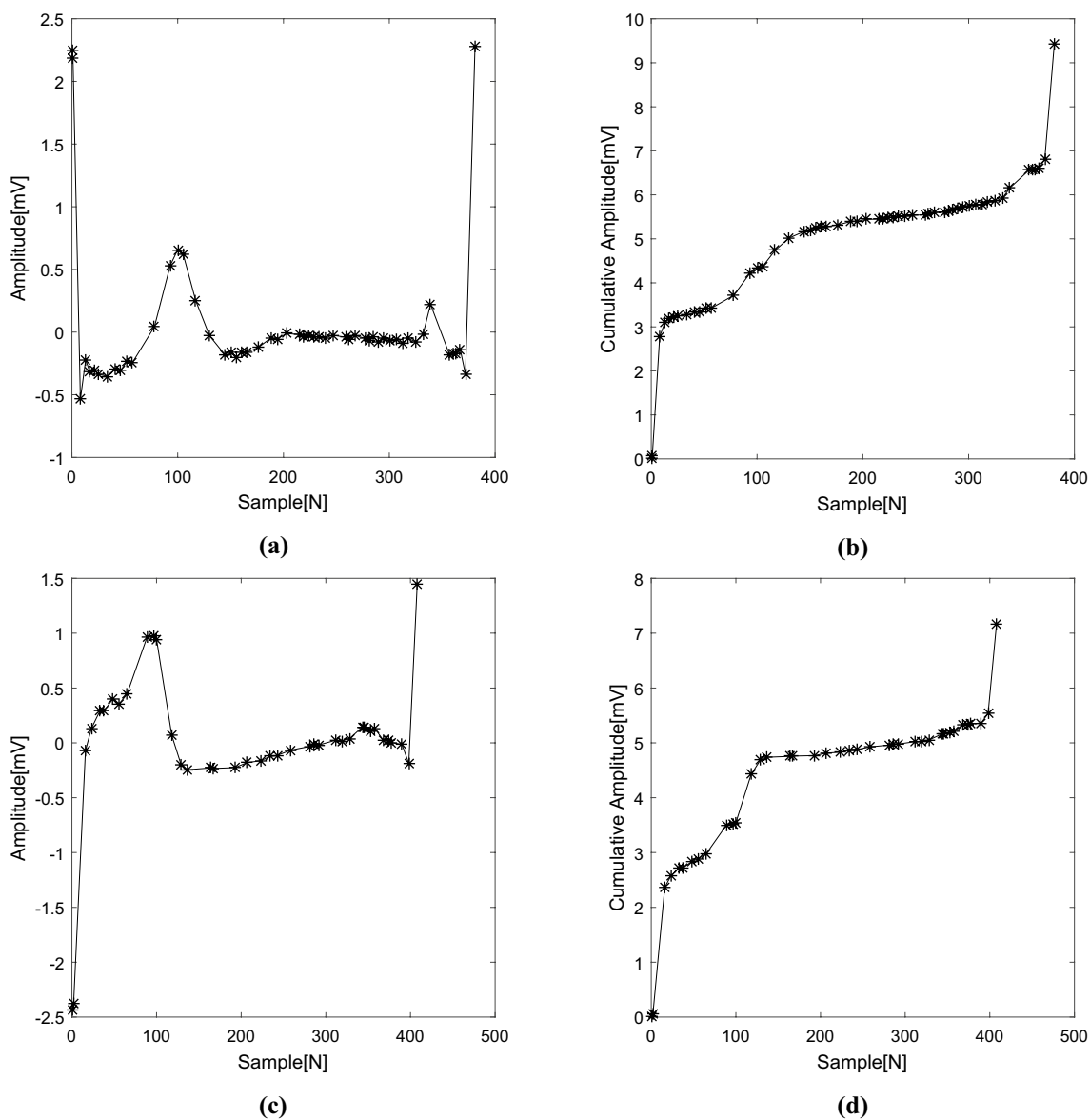


Fig. 10 Comparison of cumulative signal results according to waveform type: **a** upward QRS complex, **b** cumulative signal of **a**, **c** downward QRS complex, and **d** cumulative signal of **(d)**

The cumulative signal monotonically increases, even if the QRS complex appears as a downward wave or includes Q- and S-peaks that appear as downward or upward waves (Fig. 10).

3.9 FP detection

The onset and offset are the boundaries between the waveform and baseline. The amplitude is low in the baseline region, and the amplitude change rapidly occurs in the waveform region. Thus, the FP in vertices has the following features shown in Fig. 11.

After obtaining the feature values of each vertex, amplitude difference (A_i), time difference (T_i), and angles with neighboring vertices (θ_{i_L} and θ_{i_R}) (Fig. 11), the vertex with the highest probability of being an FP is determined as follows:

$$v_{Q_{on}} = argmax \left\{ \frac{A_{i_R}}{\max(A_{i_R})} + \frac{T_{i_R}}{0.3} + \sqrt{\cos\theta_{i_L} + \sin^2\theta_{i_R}} \right\}. \tag{8}$$

Similarly, we can obtain the S-offset of the QRS complex. However, compared with the Q-onset, the distribution of the S-offset is unclear due to the influence of

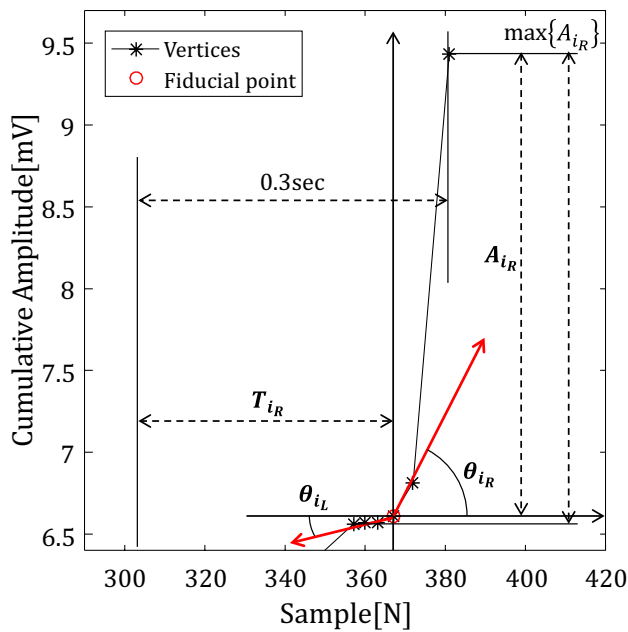


Fig. 11 Vertex features for QRS onset

irregular T-wave, making it difficult to apply the proposed method. Therefore, in this study, based on the Q-onset detection result, the QR interval, which is the distance from the R-peak to Q-onset, is used instead of the QRS interval to detect abnormal beats.

4 Proposed method

The proposed method consists of three steps. First, to minimize the number of bits of vertices' time information, a method of storing the time difference with the previous vertex is introduced. Then, after detecting the FP using (8), we analyze the distribution of the feature values of the normal and abnormal beats. Since the normal beats are more and form a normal distribution and the abnormal beats are out of the distribution of normal beat, we determined the threshold values for each feature and detected the abnormal beats.

4.1 Signal compression

Linear approximation of the ECG signal can effectively represent a signal with a small number of vertices. Each vertex of an approximated signal consists of the time and amplitude information. The amplitude is the same as the input signal because the vertex is determined among the existing samples. However, compared with the periodic input signal, the approximated signal comprises aperiodic vertices. Therefore, the approximated signal requires additional storage for time information of the vertex. In addition, the time information requires more storage as the sampling frequency and measurement time increase. Accordingly, if there is no postprocessing for additional time information, there is a problem that the amount of the compressed signal is further increased due to the storage for the added time information.

In this study, to minimize the amount of time information of the vertices, the constraint on the time difference between vertices in Sect. 3.4 is used. We express the time information

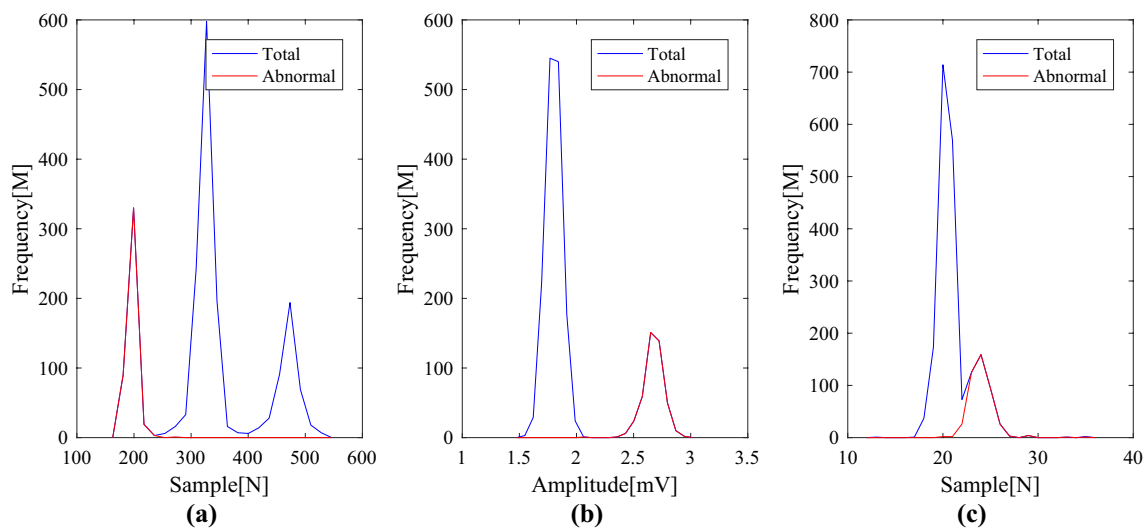


Fig. 12 Histogram of Datum 119's features: **a** RR interval, **b** amplitude of R-peak, and **c** QR interval

Table 1 Signal compression ratio of linear approximation

Raw signal	Sample × amplitude 648,964 × 12 = 7,787,568 bit
Linear approximation	
Approximation of beats	Vertices × (vertex interval + vertex amplitude) 65,008 × (5 + 12) = Total 1,105,136 bit
Compression ratio	100% × (1 - 1,235,216/9,085,496) = 85.81% 7.05: 1

of the i th vertex (T_i) by accumulating the time difference from the previous vertex (T'_i) as follows:

$$T_i = \sum_{k=1}^i T'_k, \tag{9}$$

where $T'_1 = T_1$ and $T'_i = T_i - T_{i-1}$.

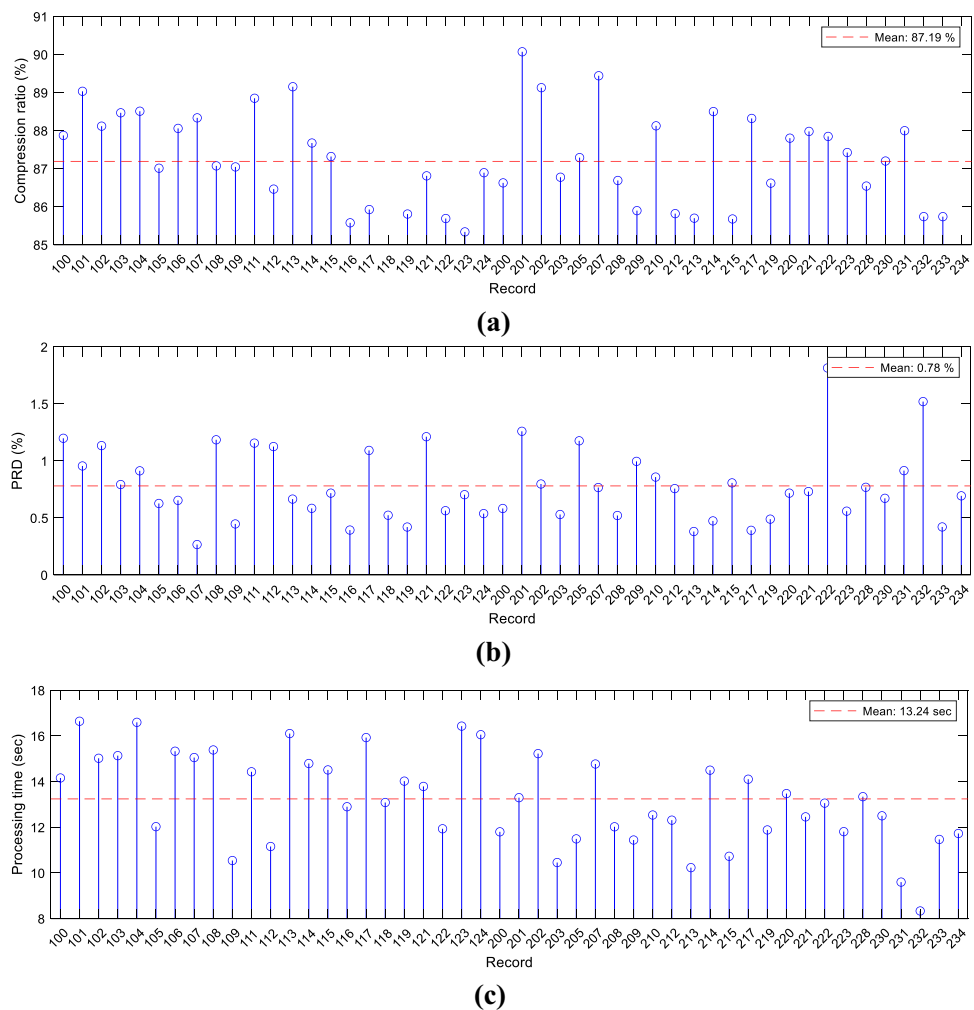
The compression ratio (CR) of an approximated signal is calculated by the ratio between the total bits of the input signal's amplitude and that of the approximated vertices' amplitude and time difference as follows:

$$CR = \left(1 - \frac{\text{Total bits of vertices}}{\text{Total bits of signal}} \right) \times 100\% = \left(1 - \frac{V_N \times (\text{Bit}_A + \text{Bit}_T)}{L \times \text{Bit}_A} \right) \times 100\% . \tag{10}$$

4.2 Feature extraction

From the approximated signal, we obtain the onset and offset of the QRS complex. Generally, the distribution of feature

Fig. 13 Distribution of experimental results for MIT-BIH ADB records: **a** compression ratio, **b** PRD, and **c** processing time



values of normal beats is similar to the normal distribution, and the distribution of abnormal beats deviates from the center of the normal beat distribution. Figure 12 shows the distribution of feature values for Datum 119, where PVC abnormal beats occurred.

The distribution of the feature values of the normal beat is similar to the normal distribution, and the distribution of the characteristic values of the abnormal beats deviates from the normal beat distribution (Fig. 12).

4.3 Abnormal beat detection

In this study, we detect PVC, which is a typical abnormal beat. As shown in Fig. 12, we can detect PVCs using various representative features—RR interval (RR), amplitude (A) of R-peak, and QR interval (QR). For PVCs, the RR interval is shorter than the normal beats, the amplitude is larger than the normal beats, and the QR interval is longer than the normal beats.

In the abnormal beat detection, the i_{th} beat is determined to be an abnormal beat if it satisfies any of the following conditions:

$$\begin{cases} RR_i \leq RR_\mu - RR_\sigma * RR_k \\ |A_i - A_\mu| \geq +A_\sigma * A_k \\ QR_i \geq QR_\mu + QR_\sigma * QR_k, \end{cases} \quad (11)$$

where μ , σ , and k are the average, standard deviation, and weight of the feature values, respectively.

5 Experiment

We experimented using the Massachusetts Institute of Technology-Beth Israel Hospital Arrhythmia Database (MIT-BIH ADB) that Physionet provided (Moddy and Mark 1990). The MIT-BIH ADB comprised 48 records, including various types of abnormal beats. Each record was 30 min long with a 360-Hz sampling frequency. We confirmed the performance of the signal compression and then confirmed the performance of abnormal beat detection. The experiments were conducted on Window 10 64-bit, Intel i5-10400, 2.90-GHz CPU, 32-GB DDR4 RAM, and MATLAB R2021a.

5.1 Signal compression

Each record of MIT-BIH ADB comprised approximately 650,000 samples, and each sample contains amplitude information (12 bits). In addition, each vertex contained

amplitude information (12 bits) and the time difference from the previous vertex (5 bits).

Table 1 shows an example of a detailed signal compression ratio of Datum 119.

The approximation error for the input signal (X) and the approximated signal (Y) with signal length (L) were calculated using percentage root-mean-square difference (PRD), as follows:

$$PRD(X, Y) = 100\% \times \sqrt{\frac{\sum_{i=1}^L (X(i) - Y(i))^2}{\sum_{i=1}^L X^2(i)}}. \quad (12)$$

Figure 13 shows the distribution of compression ratio, PRD, and execution time of linear approximation for MIT-BIH ADB's records.

The mean of the compression ratio was 87.19% (7.05:1), which confirmed excellent performance. Generally, the quality of PRD is rated as 0–2%: very good and 2–9%: good (Zigel et al. 2000). Datum 222, the 41st datum in Fig. 13, had the worst PRD (1.82%) but was classified as very good. The average processing time of the linear approximation was approximately 13.24 s. Considering that the input signal was acquired at approximately 30 min in length with 2300 beats, real-time processing was possible with an execution time of approximately 5.8 ms/beat.

5.2 Abnormal beat detection

We detected the onset of the QRS complex from the linear approximated signal using (8). Then, we obtained the QR interval from the Q-onset to R-peak.

Figure 14 provides an example of the results of the proposed method.

To confirm the performance of FP detection from the linear approximated signal, we confirmed the PVC detection performance using (11). First, we confirmed the detection result using only the QR interval feature. Then, we compared the result using the R-peak's amplitude and RR interval features with the result when the QR interval feature was additionally used. Considering that the QR interval was used, we experimented on records, which contained the abnormal beat that had a wider QR interval than the normal beat.

We evaluated classification performance by measuring sensitivity (Se) and specificity (Sp) according to true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP and TN were the correct classification results of abnormal and normal beats, respectively. Meanwhile, FP and FN were the misclassification results for normal and abnormal beats, respectively. Se , Sp , and accuracy (Ac) measure the nondetection, overdetection, and correct total beat detection rates, respectively, as follows:

Fig. 14 Example of results of the proposed method for Datum 119

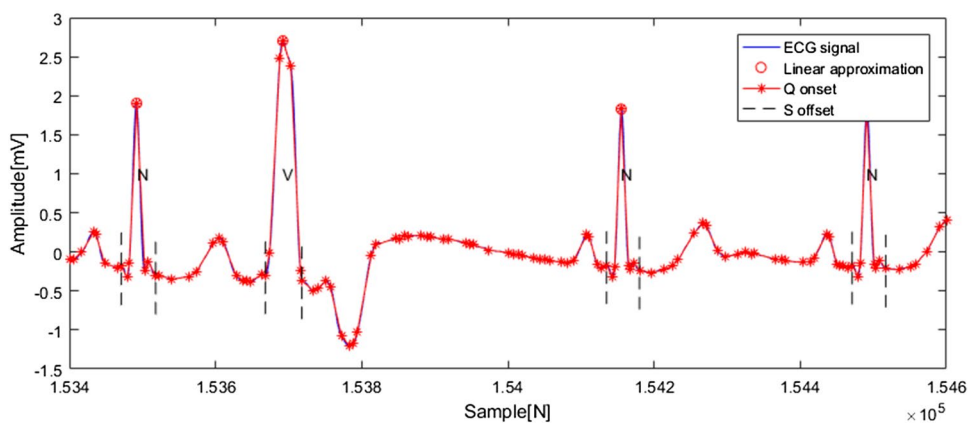


Table 2 Abnormal beat detection result using the QR interval

Record	Total beat	PVC beat	TP	TN	FP	FN	Se (%)	Sp (%)	Ac (%)
106	2023	519	398	1413	91	121	76.69	93.95	89.52
119	1984	443	413	1539	1	30	93.23	99.94	98.44
200	2565	825	629	1634	106	196	76.24	93.91	88.23
208	2575	991	788	1555	29	203	79.52	98.17	90.99
221	2423	396	388	2025	2	8	97.98	99.90	99.59
233	3057	830	750	2173	54	80	90.36	97.58	95.62
Total	14,627	4004	3366	10,339	283	638	84.07	97.34	93.70

Table 3 Abnormal beat detection result using R-peak's amplitude and RR interval

Record	Total beat	PVC beat	TP	TN	FP	FN	Se (%)	Sp (%)	Ac (%)
106	2023	519	519	1474	30	0	100.00	98.01	98.52
119	1984	443	443	1533	7	0	100.00	99.55	99.65
200	2565	825	803	1733	7	22	97.33	99.60	98.87
208	2575	991	394	1353	231	597	39.76	85.42	67.84
221	2423	396	389	1800	227	7	98.23	88.80	90.34
233	3057	830	678	2226	1	152	81.69	99.96	95.00
Total	14,627	4004	3226	10,119	503	778	80.57	95.26	91.24

Table 4 Abnormal beat detection result using R-peak's amplitude, RR interval, and QR interval

Record	Total beat	PVC beat	TP	TN	FP	FN	Se (%)	Sp (%)	Ac (%)
106	2023	519	519	1383	121	0	100.00	91.95	94.02
119	1984	443	443	1532	8	0	100.00	99.48	99.60
200	2565	825	822	1627	113	3	99.64	93.51	95.48
208	2575	991	874	1332	252	117	88.19	84.09	85.67
221	2423	396	396	1798	229	0	100.00	88.70	90.55
233	3057	830	816	2172	55	14	98.31	97.53	97.74
Total	14,627	4004	3870	9844	778	134	96.65	92.68	93.76

$$Se = \frac{TP}{TP + FN}, Sp = \frac{TN}{TN + FP}, Ac = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

Table 2 shows the abnormal beat detection result using the QRS interval.

As shown in Table 2, we can detect many PVCs from the approximated signals using only the QR interval feature.

Table 3 shows the abnormal beat detection performance for the records in Table 2 when the R-peak's amplitude and RR interval are used instead of the QR interval.

Table 4 shows the abnormal beat detection performance when the QR interval is additionally used.

Comparing the results of Tables 3 and 4, we confirmed that the overall detection result was improved. In particular, the non-detection that affected false diagnosis was significantly reduced.

Non-detected and over-detected beats occur due to ambiguous feature distribution between normal and abnormal. In addition, in some beats, error in feature values occurs due to erroneous detection of FPs. These misdetections occur when the FP does not appear well as a vertex because the number of vertices used for approximation is insufficient, or, conversely, when the number of vertices is excessive and neighboring vertices are not acquired correct feature values.

Therefore, in future research, it is necessary to research about reliable determination of the number of vertices suitable for detecting the FP, and to further improve the method for determining neighboring vertices to minimize distortion caused by noise or baseline movements.

6 Conclusion

In this article, we demonstrated effective signal compression and abnormal beat detection based on a linear approximation. Using the constraint of the time difference between vertices used in the existing linear approximation method, we improved the signal compression rate. Experimentally, we confirmed the excellent performance of the compression rate, PRD, and execution time of the proposed signal compression. In addition, by analyzing the vertex of the approximated signal to detect the FP of the QRS complex and detecting PVC abnormal beats using the width of the QR interval, we confirmed that the proposed method enabled the abnormal beat detection without signal reconstruction. Therefore, we expect the proposed method to be effective in signal analysis in low-memory embedded devices due to the characteristic of detecting abnormal beats without compressed signal reconstruction.

In future studies, by additionally detecting the FPs of the P- and T-wave from approximated signals, we will extend the types of abnormal beat that can be detected from the linear approximated signal. Further, we will modify the proposed method for determining the number of vertices, which is determined regardless of the compression ratio or reconstruction error, to improve the efficiency of signal

compression and emphasize FP features rather than the surrounding vertices.

Acknowledgements This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea Government (MSIT), Metamorphic Approach of Unstructured Validation/Verification for Analyzing Binary Code, under Grant 2021-0-00944, 60%, in part by the Ministry of Education through the BK21 FOUR Project under Grant 4199990113966, 10%, in part by the National Research Foundation of Korea (NRF) funded by the Ministry of Education through the Basic Science Research Program under Grant NRF-2020R111A1A01072343, 20%, and Grant NRF-2018R1A6A1A03025109, 10%.

References

- Arshad A, Khan S, Alam AHMZ, Tasnim R, Boby RI (2016) Health and wellness monitoring of elderly people using intelligent sensing technique. In: 2016 international conference on computer and communication engineering (ICCCCE), pp 231–235. <https://doi.org/10.1109/ICCCCE.2016.58>
- Bellman R, Dreyfus S (2015) Applied dynamic programming. Princeton Legacy Library, Princeton University Press, Princeton
- Bera P, Gupta R, Saha J (2020) Preserving abnormal beat morphology in long-term ECG recording: an efficient hybrid compression approach. *IEEE Trans Instrum Meas* 69:2084–2092. <https://doi.org/10.1109/TIM.2019.2922054>
- Huszar RJ (2007) Basic dysrhythmias: interpretation and management. Mosby Jems/Elsevier, Maryland Heights
- Illanes-Manriquez A, Zhang Q (2008) An algorithm for robust detection of QRS onset and offset in ECG signals. In: 2008 computers in cardiology, pp 857–860. <https://doi.org/10.1109/CIC.2008.4749177>
- James AP (2015) Heart rate monitoring using human speech spectral features. *Hum Centric Comput Inf Sci* 5:1–12. <https://doi.org/10.1186/s13673-015-0052-z>
- Laguna P, Jané R, Caminal P (1994) Automatic detection of wave boundaries in multilead ecg signals: validation with the cse database. *Comput Biomed Res* 27:45–60. <https://doi.org/10.1006/cbmr.1994.1006>
- Lee S, Park D (2020) Enhanced dynamic programming for polygonal approximation of ECG signals. In: 2020 IEEE 2nd global conference on life sciences and technologies (LifeTech), pp 121–122. <https://doi.org/10.1109/LifeTech48969.2020.1570620076>
- Lee S, Park D (2021a) Improved dynamic programming in local linear approximation based on a template in a lightweight ECG signal-processing edge device. *J Inf Process Syst* 18(1):97–114. <https://doi.org/10.3745/JIPS.03.0173>
- Lee S, Park D (2021b) A real-time abnormal beat detection method using template cluster for ECG diagnosis on IoT devices. *Hum Cent Comput Inf Sci*. <https://doi.org/10.22967/HGIS.2021.11.004>
- Lee S, Park D, Park KH (2017) QRS complex detection based on primitive. *J Commun Netw* 19:442–450. <https://doi.org/10.1109/JCN.2017.000076>
- Lee S, Jeong Y, Park D, Yun BJ, Park KH (2018) Efficient fiducial point detection of ECG QRS complex based on polygonal approximation. *Sensors* 18:1–16. <https://doi.org/10.3390/s18124502>
- Lee W, Kim N, Lee B (2019a) An adaptive transmission power control algorithm for wearable healthcare systems based on variations in the body conditions. *J Inf Process Syst* 15(3):593–603. <https://doi.org/10.3745/JIPS.03.0118>

- Lee S, Jeong Y, Kwak J, Park D, Park KH (2019b) Efficient communication overhead reduction using polygonal approximation-based ECG signal compression. In: 2019b international conference on artificial intelligence in information and communication (ICAIIIC), pp 058–061. <https://doi.org/10.1109/ICAIIIC.2019b.8668974>
- Lee S, Jeong Y, Kwak J, Park D, Park KH (2019c) Advanced real-time dynamic programming in the polygonal approximation of ECG signals for a lightweight embedded device. *IEEE Access* 7:1628501–2162861. <https://doi.org/10.1109/ACCESS.2019.2952399>
- Li A, Wang S, Zheng H, Ji L, Wu J (2010) A novel abnormal ECG beats detection method. In: 2010 the 2nd international conference on computer and automation engineering (ICCAE), pp 47–51. <https://doi.org/10.1109/ICCAE.2010.5452002>
- Lin HY, Liang SY, Ho YL, Lin YH, Ma HP (2014) Discrete-wavelet-transform-based noise removal and feature extraction for ECG signals. *IRBM* 35:351–361. <https://doi.org/10.1016/j.irbm.2014.10.004>
- Madeiro JP, Cortez PC, Marques JA, Seisdedos CR, Sobrinho CR (2012) An innovative approach of QRS segmentation based on first-derivative, hilbert and wavelet transforms. *Med Eng Phys* 34:1236–1246. <https://doi.org/10.1016/j.medengphy.2011.12.011>
- Manriquez AI, Zhang Q (2007) An algorithm for QRS onset and offset detection in single lead electrocardiogram records. *Annu Int Conf IEEE Eng Med Biol Soc*. <https://doi.org/10.1109/IEMBS.2007.4352347>
- Martinez A, Alcaraz R, Rieta JJ (2010) Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. *Physiol Meas* 31:1467–1485. <https://doi.org/10.1088/0967-3334/31/11/005>
- Meng Y, Yi S, Kim H (2019) Health and wellness monitoring using intelligent sensing technique. *J Inf Process Syst* 15(3):478–491. <https://doi.org/10.3745/JIPS.04.0115>
- Moddy GB, Mark RG (1990) The MIT-BIH arrhythmia database on CD-ROM and software for use with it. In: *Proceedings computers in cardiology*, pp 185–188. <https://doi.org/10.1109/CIC.1990.144205>
- Mohd Apandi ZF, Ikeura R, Hayakawa S, Tsutsumi S (2020) An analysis of the effects of noisy electrocardiogram signal on heartbeat detection performance. *Bioengineering* 7(2):53. <https://doi.org/10.3390/bioengineering7020053>
- Mokhtarian F, Suomela R (1998) Robust image corner detection through curvature scale space. *IEEE Trans Pattern Anal Mach Intell* 20:1376–1381. <https://doi.org/10.1109/34.735812>
- O’Connell KJ (1997) Object-adaptive vertex-based shape coding method. *IEEE Trans Circuits Syst Video Technol* 7:251–255. <https://doi.org/10.1109/76.554440>
- Ou Y, Li X, Guo Z, Wang Y (2020) Anobeat: anomaly detection for electrocardiography beat signals. In: 2020 IEEE fifth international conference on data science in cyberspace (DSC), pp 142–149. <https://doi.org/10.1109/DSC50466.2020.00029>
- Ozdemir MA, Guren O, Cura OK, Akan A, Onan A (2020) Abnormal ECG beat detection based on convolutional neural networks. In: 2020 medical technologies congress (TIPTEKNO), pp 1–4. <https://doi.org/10.1109/TIPTEKNO50054.2020.9299260>
- Pan J, Tompkins WJ (1985) A real-time QRS detection algorithm. *IEEE Trans Biomed Eng BME-32*:230–236. <https://doi.org/10.1109/TBME.1985.325532>
- Rhim H, Tamine K, Abassi R et al (2018) A multi-hop graph-based approach for an energy-efficient routing protocol in wireless sensor networks. *Hum Cent Comput Inf Sci*. <https://doi.org/10.1186/s13673-018-0153-6>
- Teraoka T (2012) Organization and exploration of heterogeneous personal data collected in daily life. *Hum Cent Comput Inf Sci*. <https://doi.org/10.1186/2192-1962-2-1>
- Yazdani S, Vesin JM (2016) Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology. *Digit Signal Process* 56:100–109. <https://doi.org/10.1016/j.dsp.2016.06.010>
- Zigel Y, Cohen A, Katz A (2000) The weighted diagnostic distortion (WDD) measure for ECG signal compression. *IEEE Trans Biomed Eng* 47(11):1422–1430. <https://doi.org/10.1109/TBME.2000.880093>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.