Understanding Bit-Error Trade-off of Transform-based Lossy Compression on Electrocardiogram Signals

Aekyeung Moon
ETRI
akmoon@etri.re.kr

Seung Woo Son
UMass Lowell
seungwoo_son@uml.edu

Jiuk Jung
KIMM
jijoung@kimm.re.kr

Yun Jeong Song
ETRI
yjsong@etri.re.kr

***Abstract*—The growing demand for recording longer ECG signals to improve the effectiveness of IoT-enabled remote clinical healthcare is contributing large amounts of ECG data. While lossy compression techniques have shown potential in significantly lowering the amount of data, investigation on how to trade-off between data reduction and data fidelity on ECG data received relatively less attention. This paper gives insight into the power of lossy compression to ECG signals by balancing between data quality and compression ratio. We evaluate the performance of transformed-based lossy compressions on the ECG datasets collected from the Biosemi ActiveTwo devices. Our experimental results indicate that ECG data exhibit high energy compaction property through transformations like DCT and DWT, thus could improve compression ratios significantly without hurting data fidelity much. More importantly, we evaluate the effect of lossy compression on ECG signals by validating the R-peak in the QRS complex. Our method can obtain low error rates measured in PRD (as low as 0.3) and PSNR (up to 67) using only 5% of the transform coefficients. Therefore, R-peaks in the reconstructed ECG signals are almost identical to ones in the original signals, thus facilitating extended ECG monitoring.**

***Index Terms*—Transform coding, Lossy compression, IoT, Health care, R-peak, data fidelity**

I. Introduction

Various application domains benefit from the Internet of Things (IoT) that enables the extraction of suggestive knowledge from plenty of raw datasets continuously collected by IoT devices. One of those IoT-based services of considerable importance is remote healthcare applications, where IoT enables medical monitoring, diagnosis, and treatment services to be delivered efficiently through a digital medium such as mobile devices. A recent study reports that those applications have the potential to generate about $1.1-$2.5 trillion of annual economic impact globally by 2025 [1]. As cardiovascular diseases (CVDs) are the number one cause of death globally, medical researchers have placed significant importance on electrocardiogram (ECG) analysis as it is one of the most common methods for testing clinical cardiac functions [2].

Current medical screening and diagnostic procedures, which traditionally handle the majority of data processing on remote systems that have higher computational power and storage capacity [3], have shifted toward storing longer ECG signals with the convergence of AI technologies and increasing computational power. However, in mobile healthcare services, the volume of ECG signals generated has increased at an unprecedented rate, making it practically infeasible to send all raw data to the remote site for processing. One potential solution to this is to apply compression, particularly lossy, techniques such that the storage and communication overheads are reduced [4]–[6]. As lossy compression uses inexact approximations, understanding the impact of data compression techniques in terms of bit and error rate could be of significance in making ECG records broadly usable in remote clinical healthcare.

There are a variety of data compression algorithms proposed to meet different application needs. However, prior ECG compression methods in the literature have reported limited discussion on the performance of the compressed and the reconstructed ECG signals in terms of the R-peak detection accuracy [3], [7]. In this paper, we first evaluate the impact of lossy compression algorithms on ECG signals based on two well-known transformations, namely, discrete cosine transform (DCT) [8] and discrete wavelet transform (DWT). To evaluate the impact on data fidelity, we also validate the value and the position of R-peaks in ECG signals between original datasets and reconstructed ones. Our results demonstrate that R-peak accuracy did not drop much even if data is reconstructed using only 3% of coefficients, which matches 99.9% of energy (or information). In most cases, the detected R-peaks in the reconstructed data are almost equivalent to the ones based on the original data. These results indicate that lossy compression has a beneficial effect not only on reducing ECG data significantly to minimize the burden on storage and transmission but also on maintaining high data quality.

Overall our experimental results demonstrate: 1) our DCT-based lossy algorithms generate competitive compression ratios (83%-92%); 2) the error introduced by our lossy compressions is marginal in terms of measured distortion level in PRD (percentage root-mean-square difference) and PSNR (peak signal-to-noise ratio) while using only 5% of the original signal; and 3) the reconstructed R-peaks almost coincide with ones in the original signal.

II. Background

A. IoT Services for ECG

The extended measure of ECG signals, a useful diagnostic signal to evaluate the electrical and muscular functions of

978-1-7281-6251-5/20/$31.00 ©2020 IEEE

the heart, during ordinary daily activities serves to detect and characterize anomaly in cardiac functions [9]. Thus, developing a reliable and scalable IoT-enabled ECG system is essential for accurate and efficient remote screening and diagnosis [3]. At the same time, IoT devices should minimize the use of limited storage, computing, and bandwidth [10]. Therefore, leveraging the understanding of the impact of lossy compression on ECG data in terms of bits and errors presented in this study can facilitate remote patient treatment under edge computing environments by connecting patient networks with a medical infrastructure.

B. Discussion of Prior Work

While the lossless compression method recovers the original ECG signal precisely from the compressed signal, lossy compression methods discard (or eliminate) some ECG samples from the original datasets to achieve much higher compression ratios. Several lossy compression methods for ECG data in the literature [11] reported that the eliminated data may or may not be noticeable to the user depending on the type of lossy compression employed. Thus in the context of applying lossy compression techniques [12], it is vital not only to reduce the data size significantly but also to quantify error rates, the latter of which might incur the risk of potential distortion.

As the medical signals contain and relay essential information required for precise detection of diseases, any noticeable distortion can result in an erroneous diagnosis, which could turn out to be a fatal consequence. However, for fast and efficient transmission of a large amount of information, it might be acceptable to overlook a certain amount of distortion as long as it will not hurt data fidelity much. However, achieving a lossy compression technique with a lesser amount of distortion in biomedical data is a challenging task [13]. For example, Roy et al. compared the recovered signals using various levels of PSNR. However, they did not consider features such as R-peak. Rebollo-Neira [9] also introduced the effectiveness of the transform-based lossy compression and distortion using PRD but did not deal with the characteristics of ECG directly.

In our recent studies [14], [15], we showed that transformation-based lossy compressions could ensure that reconstruction errors are within a tolerable range for datasets from IoT enabled farm and HPC applications. No prior works, however, have studied the effect of lossy data compression on the fidelity of ECG signals. For the validation of data fidelity from the reconstruction of compressed datasets, [16] evaluated microclimate datasets collected in smart farm domains and obtained the promising results that classification models can effectively predict using the reconstructed datasets. [3] proposed several compression methods and compared each of them in terms of QRS detection accuracy.

C. ECG Signals Processing

We use the ECG datasets collected using ActiveTwo by Biosemi, as depicted in Figure 1a. The ECG datasets include 2,048 data points per second. Figure 1b shows the results after applying filters from 0.3 Hz (using low pass filter) to 35 Hz (using high pass filter). After filtering, the signal on Figure 1b shows an internal characteristics representation of ECG, such as R-peak in the PQRST [17] complex [7].

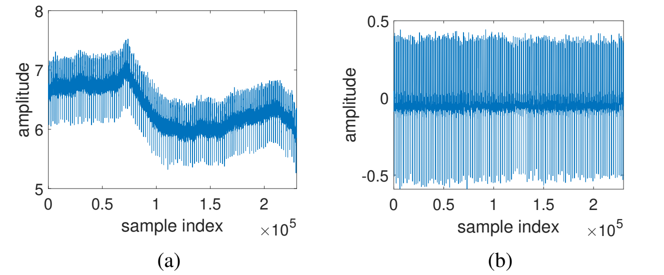


Fig. 1. ECG transformation (a) Original ECG signals. (b) after filtering (0.3 Hz by low pass filter), 35Hz by high pass filter).

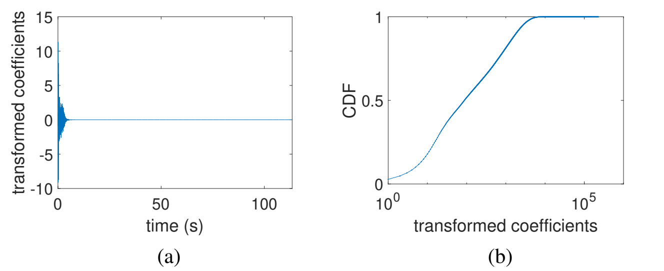


Fig. 2. The transformed representations of ECG signals. (a) after applying DCT transform. (b) Cumulative distribution function (CDF) in the sequences of transformed DCT coefficients.

In transform-based lossy compression algorithms [18], unlike original signals, the transformed signals often contain *only* a few significant components [15], [19], [20]. In other words, discrete transforms redistribute the energy contained in the signal and concentrate it into a small number of dominant coefficients (as low-frequency coefficients).

To demonstrate that similar signal property of high energy compaction exhibits in ECG, Figure 2 shows the distribution of coefficients (in the time domain) after applying a discrete transform (DCT in this case). Figure 2 depicts a case where DCT components containing 99%-percentile of the energy attained by the original data are only 2% (4,667 out of 232,400) of the entire ECG datasets (shown in Figure 1b). In this paper, we use the energy spectrum of signal theory, where a given signal is the sum of the squared transform coefficients [8], [15]. Figure 2b illustrates such a high correlation through the energy cumulative distribution function (CDF) for DCT coefficients. The x-axis is represented as a log scale to emphasize that a small number of transformed components account for most of the energy.

III. Quantifying Effects of Lossy Compression

A. ECG Data Analytics Requirements

In typical use cases, server systems carry out processing ECG signals collected for an extended period as they have higher memory/storage capacity and high-performance processors. Maintaining the same level of data processing, however, is hardly feasible for edge devices, which in general have limited storage, processor, and processing power.

As far as handling ECG signals is concerned, there are three use case scenarios. The first scenario is collecting ECG signals for offline analysis and then process them in the server later. Secondly, ECG signals are collected and analyzed in real-time at the device. For example, Gradl et al. [21] considered Android-based mobile devices to allow real-time ECG monitoring. The last scenario is collecting ECG signals in real-time, transmit signals to remote servers or cloud, and investigate the transmitted signals there. The compression framework discussed in this paper could be beneficial to the third case, as our compression scheme could compress data with high fidelity at minimum bit rates.

Our motivations in evaluating transform-based lossy compression techniques on ECG are as follows. First, applying lossy compression based discrete transforms such as DCT and DWT gives a reasonably high compression performance because such transforms can easily decorrelate inherent randomness exhibited in ECG signals. Second, although individual data values in ECG signals show a certain degree of randomness, their overall patterns are smooth spatiotemporally. Because of this, combining data transformation with compression can be more effective as the transformed data usually reveal the correlation of the data explicitly. Lastly, R-peak is one of the most critical features in the ECG signal, which plays a vital role in the diagnosis of heart rhythm abnormalities and also in determining heart rate variability (HRV) [22], [23]. Therefore, we validate the fidelity of reconstructed datasets based on R-peak variations.

B. Data Compression

Our compression approach begins by transforming raw ECG data into sparse representations using a discrete transform, either DCT or DWT. To illustrate our mechanism in detail, let us consider $\hat{x}$, which expresses the transformed components of the original datasets (x) using either DCT [8] or DWT [24]. To model the correlation between the transformed coefficients and energy (or information) represented among them, let us further define $\hat{x}_t$, which denotes transformed components at compressed block size of n at given period t: $\hat{x}_t = \{\hat{x}_{t,1}, \hat{x}_{t,2}, ...\hat{x}_{t,n}\}$. Thus, each coefficient component has own energy coefficient defined as: $e(\hat{x}_{t,i})$. We formulate $EC(\hat{x}_{t,k})$ as the energy concentration (EC) contained in the number of coefficients components, denoted as k, of the entire transformed components ($\hat{x}_t$), which is calculated as:

$$EC(\hat{x}_{t,k}) = \frac{\sum_{n=1}^{k} e(\hat{x}_{t,n})^2}{\sum_{n=1}^{N} e(\hat{x}_{t,n})^2}, n = 1, 2, ..., N, k \leq N. \quad (1)$$

The energy concentration of the entire transformed components (n) is 1.0 (or 100%-percentile), and a partial of coefficient components, for instance, k components ($\{\hat{x}_{t,1}, \hat{x}_{t,2}, ...\hat{x}_{t,k}\}$) represents a certain portion of energy concentration. Given this notion, we use the energy threshold, δ, to select the k-dominant components from the entire transformed components. To achieve higher compression ratios, we ignore the non-significant components other than the k-dominant components. In our compression method, it is possible to achieve a higher compression ratio if k, i.e., the number of dominant coefficients, is sufficiently smaller [24]–[26].

Our compression procedure based on DCT and DWT is described in more detail as follows. The procedure starts by transforming the original data, x at a given period of t, into DCT and DWT basis components to decorrelate data. In this way, the original datasets x_t of compressed block size (n) are converted into the transformed datasets, $\hat{x}_t$. After the transform, $\hat{x}_t$ including n-coefficient components are sorted in descending order. In general, the most dominant component is $\hat{x_{t,1}}$ in the transformed coefficients [27]. Finally, we determine the k largest components (using Equation 1) that amount to the required energy threshold δ. As k-dominant components are selected from the sorted version (s_t) of $\hat{x}_t$, the approximated data points (AD_t) need to include k-dominant components and their indices, the latter of which indicate the positions of the selected components in $\hat{x}_t$ for reconstructing data correctly. All non-significant components ($n - k$) are discarded.

C. Data Reconstruction

Our reconstruction algorithm using k-dominant transformed components in conjunction with their indices is as follows. Using AD_t (at period t), we generate $\hat{x}_t$ by substituting the indices of data points with zeros except for the indices of k-dominant. We then reconstruct data points x'_t, which are generated from $\hat{x}_t$ by applying inverse transformation methods on a cosine or wavelet basis.

IV. Evaluations

A. Setup

1) Datasets: We use ECG datasets collected from Biosemi's ActiveTwo system in our evaluation (as shown in Figure 3). ActiveTwo can measure electrocardiogram (ECG) signals using sampling frequency up to 16,384 Hz (152 channels). In this study, we measured data at 2,048 Hz. In other words, it collects 2,048 data points per second. After collecting all ECG records to evaluate, we implement and evaluate our compression methods using MATLAB.

2) ECG Datasets Fidelity: We compare the morphological patterns of ECG after reconstruction. The recent use of ECG analysis allows visualization of the ECG signal patterns, composed of multiple cycles that include numerous sample points [7]. The identification of these morphological patterns is a critical step in analyzing ECG signals [17], [28].

An ECG signal is composed of a successive repetition of PQRST patterns, which include P wave, QRS complex, and T wave, as shown in Figure 4a. After the P wave begins, the declining wave soon gets a downward deflection labeled as Q wave. R-peak is a sudden upright deflection. On its decline, a slight downward deflection is the S wave. In Figure 4a, we depict PQRST by aligning the R-peak as the center. R-peaks are the highest waves of the QRS regions as shown in Figure 4b. Even though every QRS complex does not contain Q, R, and S waves accurately, the Q and R wave is always

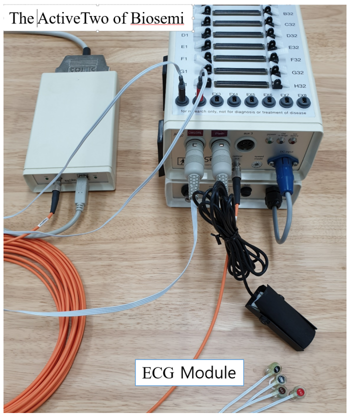
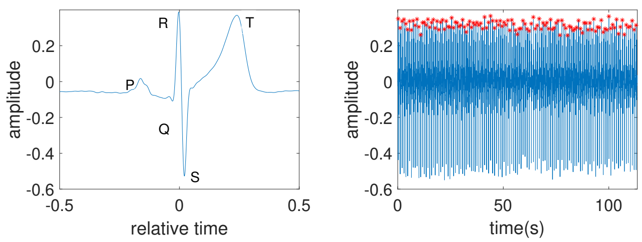


Fig. 3. Acquisition of ECG signals using Biosemi's ActiveTwo.



(a) (b)

Fig. 4. (a) PQRST wave of ECG. (b) R-peaks (in red dots).

negative and positive, respectively, in the QRS complex. These are detected by comparing relative magnitude in each QRS region. R wave is the most crucial section of this QRS, which has an essential role in the diagnosis of heart rhythm irregularities and also in determining HRV [22], [23].

3) Evaluation Metrics: We use the following metrics to assess the overall compression performance and the quality of the reconstructed data.

- Two types of compression ratio (CR), which are calculated as: $CR_K = \frac{|D|-|K|}{|D|} \times 100\%$ and $CR_A = \frac{|D|-|AD|}{|D|} \times 100\%$, where $|D|$ is the sum of x_t size, $|K|$ is the sum of k-dominant coefficients at given period t, $|AD|$ is the sum of AD_t size.
- Let x_t be the original datasets and x'_t be the reconstructed datasets at given period t to assess the quality of the reconstructed signal.
- $PSNR_t = 20log10(range) - 10log10(MSE(x_t, x'_t))$,
- $PRD_t = \frac{\|x_t - x'_t\|}{\|x_t\|} \times 100$, where $\| \ \|$ indicates the 2-norm [9].

In our performance metrics, achieving higher CR and PSNR and lower PRD is our goal.

B. Results

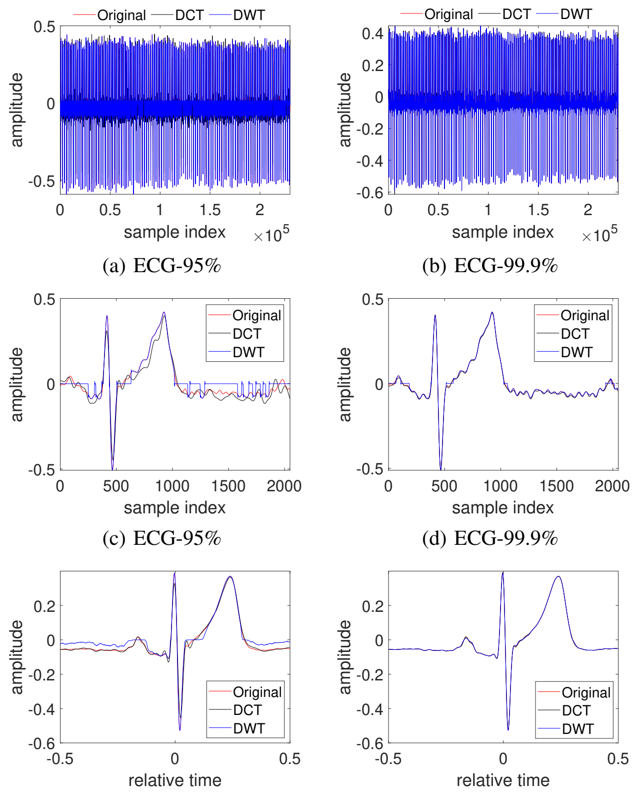
For the details about the algorithm described in Section III-B, we use the type-II DCT, which is known to generate sparse representation effectively as compared with other signal transformation methods [15]. In the case of DWT, we use Daubechnies d4 wavelet (or db4 wavelet).

1) Compression using entire datasets: We first evaluate the compression performance using the entire datasets. Table I and Figure 5 show the compression ratios when δ in Equation 1 is varied between 0.95 (95%) and 0.999 (99.9%). As shown in Table I, DCT shows higher compression ratios (CR_K and CR_A) than DWT. DWT shows lower compression ratios because of a 1-level wavelet transformation we employed, and therefore, it required increased k in Equation 1. The compression ratio of DWT increased about 1.4 times and 22.4 times in terms of CR_K and CR_A, respectively, as we decreased δ. CR_A (shown in Section III-B) refers to the actual data size being transferred because the approximated data points include k-dominant components as well as their indices. The index is required to reconstruct data correctly because it indicates the position of the selected components in $\hat{x}_t$.

Figure 5 shows the comparison of the recovered data in various cases. The error threshold of 95% overall demonstrates somewhat higher error rates than those of 99.9% (or 0.999). DWT shows higher variations than DCT, although the general trend is similar to other reconstructed data. We can also see that the reconstructed data with a small number of coefficients almost coincide with the original data. This suggests the importance of selecting compression coefficients within the tolerable error threshold for the ECG application. In the case of DCT, the effect of error varied depending on the energy threshold of δ. Specifically, the DCT and DWT show increased error rates with lower δ, as shown in Table I. The reason the error rate has similar values is that it contains different k-dominant coefficients to attain the same energy in each compression algorithm, so when reconstructed, the error appears the same according to the value of the retained energy.

The results for PSNR and PRD also showed a similar pattern. First, while DCT demonstrates better compression ratios than DWT, PSNR, and PRD values are in a similar range. In DCT, PRD in the range [0.3, 21.8] showed significant variations depending on the threshold. For a data quality measured using PRD of 0.3, the achieved average compression ratio is 87.15%. When the compression ratio increases up to 92.36%, a distortion measured in terms of PRD increases to 0.9. Like DCT, compression ratios of DWT changed significantly depending on the energy threshold, while the range of measured PRD is in [0.3, 22.3]. However, PRD of 0.3 achieved 1.98% of the compression ratio. When PRD is increased to 0.98, the compression ratio increases up to 91.4% of CR_A and 95.7% of CR_K.

2) Compression using a fixed interval: We assume that compression is performed in a fixed interval (i.e., per second) since we collected 2,048 samples per second. Figure 6 shows the variation in k values for ECG datasets per one second. As we can see, DWT requires higher k-dominant coefficients than DCT to keep the same energy. When δ = 0.95 (95%), 0.99 (99%) and 0.999 (99.9%), DCT also shows better compression ratios than DWT like the experimental results using full datasets because DCT has lower k-dominant coefficients. Like Figure 5, this case also shows that the higher the energy



(a) ECG-95% (b) ECG-99.9%

(c) ECG-95% (d) ECG-99.9%

(e) ECG-95% (f) ECG-99.9%

Fig. 5. Comparison of reconstructed datasets against the original datasets using full data points ((a) and (b)), 1 sec period or 2,048 data points ((c) and (d)), and the PQRST complex with R-peak in the center ((e) and (f)), respectively.

TABLE I
COMPARISON RESULTS USING FULL DATASETS.

Algorithm	Threshold	CR_K	CR_A	RMSE	PSNR	PRD
DCT	0.95	98.8249	97.6497	0.0315	30.0243	22.3595
	0.99	97.9918	95.9836	0.0141	37.0153	9.9980
	0.999	97.0839	94.1678	0.0045	47.0156	3.1615
	0.9999	96.3107	92.6213	0.0014	57.0146	0.9999
	0.99999	95.5581	91.1162	4.4590e-04	67.0152	0.3162
DWT	0.95	76.2169	52.4337	0.0315	30.024	22.3604
	0.99	62.7909	25.5818	0.0141	37.0138	9.9997
	0.999	55.3894	10.7788	0.0045	47.0148	3.1618
	0.9999	52.4776	4.9552	0.0014	57.0173	0.9996
	0.99999	51.1760	2.3520	4.4584e-04	67.0164	0.3161

value, the more similar the PQRST pattern to the original with increasing energy threshold.

C. Data Fidelity

In our next sets of experiments, we evaluate how lossy compression affects the quality of analytics outcomes in ECG datasets. More specifically, we assess the following features of R-peak on the reconstructed datasets.

1) Compare R-peak values in the original and the reconstructed ones.
2) Compare the time when the location of R-peak appeared in the original and the reconstructed signals.
3) Compare the index when the position of R-peak appears in the original and the reconstructed signals.

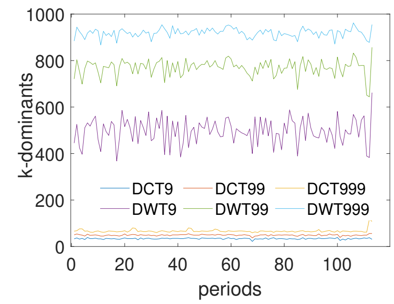


Fig. 6. Comparison of k-dominant components.

TABLE II
COMPARISON OF ERROR RATES IN R-PEAK.

Algorithm	Threshold	$D(\triangle R^s_{val})$	$D(\triangle R^s_{idx})$	$D(\triangle R^s_{tm})$
DCT	0.95	0.0519	3.1095	0.0015
	0.99	0.0117	1.3775	6.7263e-04
	0.999	0.0029	0.5542	2.7058e-04
	0.9999	6.6355e-04	0.2174	1.0613e-04
	0.99999	1.4660e-04	0.1984	9.6884e-05
DWT	0.95	0.0065	0.4695	2.2927e-04
	0.99	0.0030	0.2348	1.1463e-04
	0.999	9.0601e-04	0.1775	8.6656e-05
	0.9999	2.1153e-04	0.0887	4.3328e-05
	0.99999	5.7462e-05	0	0

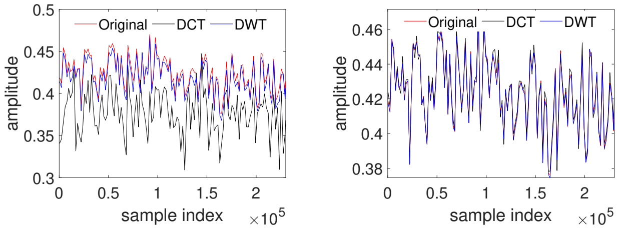
To evaluate overall data fidelity, we detect R-peaks and define the set of R-peak (R^s). R^s includes the three information about R-peak mentioned above, namely $R^s_{idx}(n)$, $R^s_{tm}(n)$, and $R^s_{val}(n)$, which measures locations, time, and value of n^{th} R-peak, respectively. Our overall evaluation metrics on the quality of R-peaks are defined as:

$$
\begin{aligned}
&R^s = \{\{R_{idx}(1), R_{tm}(1), R_{val}(1)\}, ..., \{R_{idx}(N), R_{tm}(N), R_{val}(N))\}\},\\
&\triangle R^s_{idx}(n) = x(R^s_{idx}(n) - \hat{x}(R^s_{idx}(n)),\\
&\triangle R^s_{tm}(n) = x(R_{t}ime(n) - \hat{x}(R_{tm}(n)),\\
&\triangle R^s_{val}(n) = x(R_{val}(n) - \hat{x}(R_{val}(n)).\\
&D(\triangle R^s_{idx}) = \sqrt{\frac{\sum_{n=1}^{N} \triangle R^s_{idx}(n)^2}{N}},\\
&D(\triangle R^s_{tm}) = \sqrt{\frac{\sum_{n=1}^{N} \triangle R^s_{tm}(n)^2}{N}},\\
&D(\triangle R^s_{val}) = \sqrt{\frac{\sum_{n=1}^{N} \triangle R^s_{val}(n)^2}{N}}.
\end{aligned}
$$

Figure 7 shows the comparison of R-peak values between original signals and reconstructed ones whereas Table II shows error rates in R-peaks. Overall, the energy threshold of 95% shows slightly higher deviations than that of 99.9% (or 0.999). DCT shows higher variances than DWT, although it is similar to other reconstructed data. These results suggest that we can recover data using only a small number of measurements or sampling rate. In the case of 99.99% in Table II, the position and value of R-peaks in the reconstructed data are almost identical to those in the original signal. Overall, we observe lower $\triangle R^s_{idx}(k), \triangle R^s_{tm}(k)$ and $\triangle R^s_{val}(k)$ with the higher energy threshold, which is demonstrated in Figure 7 where the reconstructed signals using 99.9% is more coincide with the original signals than those of 95%.

V. CONCLUSIONS AND FUTURE WORK

To effectively support IoT enabled remote healthcare services that produce a large volume of data, it is necessary



(a) ECG-95% (b) ECG-99.9%

Fig. 7. Comparison of reconstructed datasets against the original datasets in terms of R-peaks.

to store efficiently and transmit reliably for post-processing or diagnosis. In this paper, we evaluated ECG datasets acquired from Biosemi ActiveTwo devices as an example of a critical meta-analysis for CVD and HRV and compared the performance of lossy compression based on DCT and DWT to evaluate the feasibility of reconstructed data. Our experimental results show that lossy compression could generate significantly higher compression ratios, while the loss of data quality is acceptable. Lastly, the detected R-peaks in the reconstructed data from lossy compression techniques based on DCT and DWT are almost identical to those in original ECG signals.

We argue that lossy compression schemes can be useful for ECG datasets because the R-peak values in reconstructed data are almost identical to ones in the original data. Our experimental evaluation demonstrates that our compression method can obtain a lower distortion in terms of PRD using only 5% of the original datasets.

In our future work, we plan to experiment with other ECG datasets (Physinet), use reconstructed ECG data for real-world detection algorithms, and show how resource consumption gets affected by applying lossy compression in edge devices.

Acknowledgements

This work is supported by the Korea Innovation Foundation (INNOPOLIS) grant funded by the Korea government (MSIT) (2020-DD-UP-0278). This material is also in part based upon work supported by the National Science Foundation under Grant No. 1751143.

References

[1] J. Manyika, M. Chui, J. Bughin, R. Dobbs, P. Bisson, and A. Marrs, "Disruptive Technologies: Advances that Will Transform Life, Business, and the Global Economy," *McKinsey Global Institute*, 2013.

[2] P. E. Dilaveris, E. J. Gialafos, S. K. Sideris, A. M. Theopistou, G. K. Andrikopoulos, M. Kyriakidis, J. E. Gialafos, and P. K. Toutouzas, "Simple electrocardiographic markers for the prediction of paroxysmal idiopathic atrial fibrillation," *American Heart Journal*, pp. 733–738, 1998.

[3] M. Elgendi, A. Mohamed, and R. Ward, "Efficient ECG Compression and QRS Detection for E-Health Applications," *Scientific Reports*, 2016.

[4] T. Bose, S. Bandyopadhyay, S. Kumar, A. Bhattacharyya, and A. Pal, "Signal Characteristics on Sensor Data Compression in IoT – An Investigation," in *13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, June 2016, pp. 1–6.

[5] N. Huiibbe, A. Wegener, Y. Ling, and T. Ludwig, "Evaluating Lossy Compression on Climate Data," in *International Supercomputing Conference (ISC)*, June 2013, pp. 343–356.

[6] M. d. M. Elena, J. M. Quero, and I. Borrego, "Optimal Selection of Wavelet Coefficients for Electrocardiograph Compression," *ETRI Journal*, vol. 29, no. 4, pp. 530–532, 2007.

[7] J.-S. Park, S.-W. Lee, and U. Park, "R Peak Detection Method Using Wavelet Transform and Modified Shannon Energy Envelope," *Journal of Healthcare Engineering*, 2017.

[8] M. A. Razzaque, C. J. Bleakley, and S. Dobson, "Compression in wireless sensor networks: A survey and comparative evaluation," *ACM Transactions on Sensor Networks*, vol. 10, no. 1, p. 5, 2013.

[9] L. Rebollo-Neira, "Effective high compression of ECG signals at low level distortion," *Scientific Reports*, 2019.

[10] S. Kartakis and J. A. McCann, "Real-time Edge Analytics for Cyber Physical Systems using Compression Rates," in *11th International Conference on Autonomic Computing (ICAC '14)*, 2014, pp. 154–159.

[11] M. Abo-Zahhad, A. F. Al-Ajlouni, S. M. Ahmed, and R. Schilling, "A new algorithm for the compression of ECG signals based on mother wavelet parameterization and best-threshold levels selection," *Digital Signal Processing*, pp. 1002–1011, 2013.

[12] J. J. Chou and L. A. Piegl, "Data reduction using cubic rational B-splines," *IEEE Computer Graphics and Applications*, vol. 12, no. 3, pp. 60–68, May 1992.

[13] A. B. Roy, D. Dey, D. Banerjee, and B. Mohanty, "Comparison of FFT, DCT, DWT, WHT Compression Techniques on Electrocardiogram & Photoplethysmography Signals," in *International Conference on Computing, Communication and Sensor Network (CCSN)*, 2012.

[14] A. Moon, J. Kim, J. Zhang, and S. W. Son, "Lossy Compression on IoT Big Data by Exploiting Spatiotemporal Correlation," in *2017 IEEE HPEC*, 2017, pp. 1–7.

[15] J. Zhang, A. Moon, X. Zhuo, and S. W. Son, "Towards Improving Rate-Distortion Performance of Transform-Based Lossy Compression for HPC Datasets," in *IEEE HPEC*, 2019.

[16] A. Moon, J. Kim, J. Zhang, and S. W. Son, "Evaluating Fidelity of Lossy Compression on Spatiotemporal Data from an IoT Enabled Smart Farm," *Computers and Electronics in Agriculture*, vol. 154, Nov. 2018.

[17] F. Rundo, S. Conoci, A. Ortis, and S. Battiato, "An Advanced Bio-Inspired PhotoPlethysmoGraphy (PPG) and ECG Pattern Recognition System for Medical Assessment," in *Sensors*, 2018.

[18] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications," in *IEEE Access*, 2018, pp. 4875–4894.

[19] A. Moon, J. Kim, J. Zhang, H. Liu, and S. W. Son, "Understanding the Impact of Lossy Compression on IoT Smart Farm Analytics," in *2017 IEEE Big Data*, 2017, pp. 4602–4611.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[21] S. Gradl, P. Kugler, C. Lohmuller, and B. Eskofier, "Real-time ECG monitoring and arrhythmia detection using Android-based mobile devices," in *International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2012, p. 2452–2455.

[22] D. Sadhukhana and M. Mitra, "R-Peak Detection Algorithm for Ecg using Double Difference And RR Interval Processing," *Procedia Technology*, pp. 873–877, 2012.

[23] H. Rabbani, M. P. Mahjoob, E. Farahabadi, and A. Farahabadi, "R Peak Detection in Electrocardiogram Signal Based on an Optimal Combination of Wavelet Transform, Hilbert Transform, and Adaptive Thresholding," *Journal of Medical Signals and Sensors*, 2011.

[24] M. M. Abo-Zahhad, A. I. Hussein, and A. M. Mohamed, "Compressive Sensing Algorithms for Signal Processing Applications: A Survey," *International Journal of Communications, Network and System Sciences*, vol. 8, no. 6, pp. 197–216, 2015.

[25] R. Sustika and B. Sugiarto, "Compressive Sensing Algorithm for Data Compression on Weather Monitoring System," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, pp. 974–980, 2016.

[26] R. Chaturvedi and Y. Yadav, "A Survey on Compression Techniques for ECG Signals," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 9, 2013.

[27] J. Zhang, X. Zhuo, A. Moon, and S. W. Son, "Efficient Encoding and Reconstruction of HPC Datasets for Checkpoint/Restart," in *Symposium on Mass Storage Systems and Technologies (MSST)*, 2019.

[28] S. Yazdani and J.-M. Vesin, "Extraction of QRS fiducial points from the ECG using adaptive mathematical morphology," *Digital Signal Processing*, pp. 100–109, 2016.