2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Novel Lossless ECG Compression Algorithm for Active Implants*

Jingchuan WANG, Jin LI, Hua JIN, Xiang CHEN

***Abstract*—A low complexity lossless ECG compression algorithm for active implants is proposed in this paper. The algorithm is based on adaptive length encoding by combining linear prediction with delta encoding. The algorithm is tested on forty-eight segments of 30-min ECG signals obtained from MIT-BIH Arrhythmia Database. The results show that with the data segment length of 33 and the predictor order of 2, the average compression rate of the algorithm reaches 2.43 and there is no difference between the reconstructed signal and the original one. It implies that it can realize the lossless compression with a high compression ratio. Meanwhile, the low complexity makes this novel algorithm suitable for ECG monitoring applications of active implants.**

Keywords—ECG compression, signal processing, linear prediction, styling, compression ratio

I. INTRODUCTION

In recent years, cardiovascular disease (CVD) has been one of the leading causes of death in the worldwide, accounting for about 31% of deaths[1]. Electrocardiogram (ECG) is very crucial in the diagnosis of CVD [2] and sudden cardiac death. Holters have been widely accepted for continuous and dynamic monitoring of ECG and are very important for CVD diagnosis and prevention [3]. Although Holters can keep recording 24 to 48 hours (usually not exceeding 72 hours), it is still difficult to capture valid pathological information for certain patients. For example, only a small number of patients can record data during syncope period by using a 24-hour Holter and for the vast majority, no syncope occurred during the 24-48 hours of holters recording. It has been reported that 33% of patients with a history of Atrial Fluctuation were missed by a one-year, 24-hour ECG monitoring per month [4]. Therefore, in order to obtain valid and sufficient diagnosis information about arrhythmia and unexplained syncope, it is necessary to keep cardiac monitoring half a year or even longer.

Implantable cardiac monitors (ICM) were developed in the early 1990s for long-term ECG monitoring. [5] showed that 78% of the 218 events in 570 patients achieved ICM guided diagnosis within an average follow-up period of 10 months. However, as we know, the implantable devices are small in size and low in computing capability and power consumption. The longest service life of the commercial ICM is three years. Thus, data compression algorithms with high compression ratio and low complexity are absolutely necessary for the long-term monitoring of ICM.

At present, there are three types of compression methods: the first processes the signals in the time domain directly, such as neural network [6], or linear prediction [7]; the second in the transform domain, such as discrete cosine transform [8], or discrete wavelet packet transform [9]; the third adopts the dominant feature extraction, such as template matching [10]. The latter two methods usually have high compression ratios (CR), but they have high complexity and inevitably bring loss of information, which are not available for ICM application. The first method, in particular the linear prediction has relatively low complexity and usually is lossless, which are more suitable for ICM. Up to now, several algorithms based on linear prediction have been proposed. Delta encoding is used for a lossless ECG compression algorithm in [11]. A linear predictor and dynamic data packaging have been applied to wearable sensors [3].A lossless algorithm adopted delta predictor and Rice Golomb Coding in [12]. However, for the limitation of the ICM, more improvements are needed to obtain higher CR.

In this paper, we propose an improved ECG signal compression algorithm based on linear prediction and adaptive length encoding. The paper is organized as followings: the second section introduces the scheme and the implementation procedures of the lossless compression algorithm, the third section presents the results of the performance of the algorithm, and the conclusion is drawn in the fourth section.

II. METHODS

Generally speaking, the ECG signal of a cardiac cycle contains a QRS complex with sharp amplitude changes and other waveform segments with relatively slow changes such as P or T waves. For the latter segments, a delta encoding can eliminate effectively the redundancy. However, for the sharp QRS segments, the error signal after delta encoding is often large, resulting in low compression ratio. Actually in addition to the QRS segments, abnormal changes in ECG signal may also cause sharp amplitude changes. For this reason, additional predictors are used to reduce the prediction error of the segments with sharp changes, and obtain a higher compression ratio.

The algorithm proposed in this paper consists of compression and reconstruction. As shown in Figure 1, the

This work was supported in part by the National Natural Science Foundation of China (Grant No. 82072013, 81571761), and the Fundamental Research Funds for the Central Universities (Grant No. XJJ2015083).

J.W, J.L, X.C. is with The Key Laboratory of Biomedical Information Engineering of Ministry of Education, Institute of Health and Rehabilitation Science, School of Life Science and Technology, Xi' an Jiaotong University, Xi'an, Shaanxi, 710049, P. R. China; National Engineering Research Center for Healthcare Devices. Guangzhou, Guangdong, 510500, P.R. China; The Key Laboratory of Neuro-informatics & Rehabilitation Engineering of Ministry of Civil Affairs, Xi'an, Shaanxi, 710049, P. R. China (phone: 86-13571491940; e-mail: chenxiang@xjtu.edu.cn).

H. J. is with Lepu medical electronic instruments co., Ltd. P. R. China

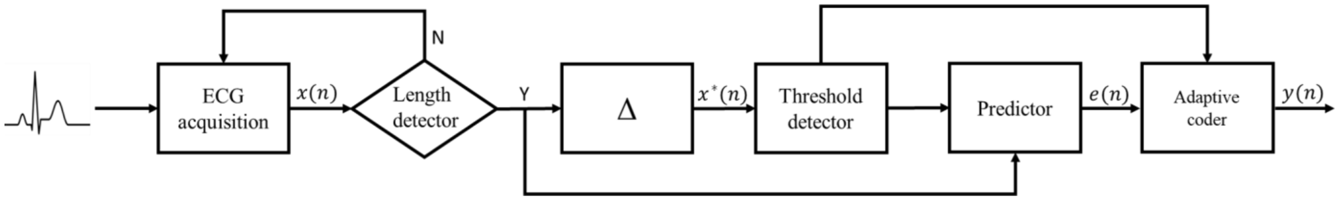

Fig 1. The block diagram of the ECG compression algorithm

4-bit	1-bit	12-bit		Min_{bit}-bit			
Min_{bit}	G_F	Header1	Header2	Error(1)	Error(2)	···	Error(N-2)

Fig 2. The typical structure of a compressed packet

compression part consists of four processing stages. The first is to transmit the real-time ECG data to the length detector, where the data are temporarily stored until it reaches the preset data segment length. The second stage is delta encoding and save the prediction error. In the third stage, the threshold detector gets the prediction error and compares it with the threshold, based on which, the prediction error is selectively transmitted to the predictor to determine the final prediction error. In the last step, the adaptive coder constructs a compressed packet with a variable length. The typical structure of a compressed packet is shown in Figure 2.

A. *Length detection*

The length detector is to set up an ECG data buffer, where the data is transferred to the delta coder when the length of the ECG data transmitted into it reaches the preset length N. The preset length N is set according to the sampling frequency of ECG signal and the number of bits of ADC. The large value of N means the large memory occupation and the long delay. Too large value of it may lead to the device cannot work normally due to the limitation of power consumption and storage space of implanted devices and the requirements for real-time processing. Therefore, it is necessary to choose an appropriate value of N.

B. *Delta Encoding*

The baseline wander is a common background noise in ECG signal, and delta encoding can eliminate it [1]. The delta encoding functions are defined as followings:

$$\hat{x}(n) = x(n-1) \quad (1)$$

$$x^*(n) = x(n) - \hat{x}(n) \quad (2)$$

where $x^*(n)$ is the predictive error after delta encoding, which is used as a criterion to determine if the data need to transfer to the additional linear predictor to further improve the compression ratio.

Figure 3 presents a segment of original ECG signal and its process result by the first-order linear predictor. It can be found that the prediction errors of the relatively slow varying waveform remain small, which indicates that the slow varying segment can be effectively predicted by the first-order linear predictor. On the contrary, the prediction errors of the waveform with sharp amplitude changes, such as QRS complex, are still large although they are obviously reduced compared with the original signal. It need to be further processed for a better performance of compression.

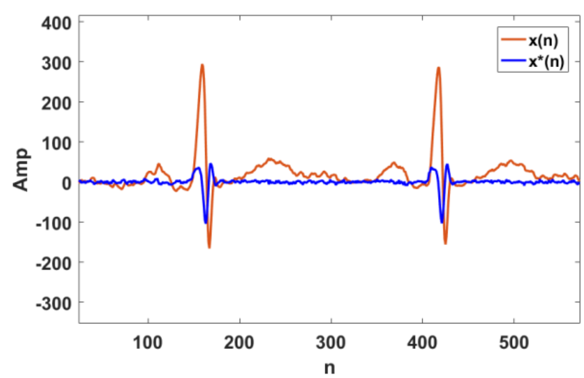

Figure 3. Comparison between the original signal and the prediction error of the first-order linear predictor

C. *Linear Prediction*

The ECG data with $x^*(n)$ above a threshold will be transferred to the linear predictor for further processing. The linear prediction functions are defined as:

$$x'(n) = \sum_{i=1}^{k} \omega(i)x(n-i) \quad (3)$$

Where $x'(n)$ is the predictive value at time t_n, $x(n-i)$ is the value of original ECG signal at time t_{n-i} and w_i represent the corresponding coefficient. The various values of order k of the linear predictor would influence the prediction effect, but they are also limited by the computational complexity. The output error $e(n)$ of the linear predictor is defined as:

$$e(n) = x(n) - \sum_{i=1}^{k} \omega(i)x(n-i) \quad (4)$$

D. Adaptive Length Encoding

The adaptive length encoding is the last step of the algorithm. The Huffman coding and the arithmetic coding are the most commonly adopted in this stage [6]. Despite of the good coding performance, the complexity of the above methods is relatively high.

A low complexity encoding method is proposed in this paper. First, it searches for the maximum value *Max* of the absolute values of the prediction errors, and calculates the minimum number of bits in accordance with equation (5):

$$2^{Min_{bit}-1} \geq Max \quad (5)$$

Min_{bit} is the number of bits required to store the value of single error. Leading zeros are filled for an insufficient length. The first bit is used as the symbol bit.

Then, the compressed package is constructed. The structure of a compressed package consists of four parts as shown in Fig.2. The first four bits store Min_{bit} calculated above. The second is the binary function G_F of threshold detection, which occupies one bit. The third is the sampling data Header of the original signal, which is stored with twelve bits. The fourth part is the prediction errors stored in sequence, each of which occupies Min_{bit} bits.

E. Signal restoration

The decompression is performed in the reverse direction of the compression. Firstly, the number of bits L of the first data segment in the compressed packet is calculated.

$$L = (G_F + 1) \times 12 + Min_{bit} \times (N - G_F - 1) \quad (6)$$

L is the binary length of the first data segment in a compressed package, and N is the preset length. The beginning of the second segment starts from L + 6 bits, and so on till the compressed packet is completely restored to multiple segments.

Next, the segments get the Headers according to the value of G_F, that is, the value is $G_F + 1$. According to the value of Min_{bit} to get the error array, and finally through a delta inverse transform to get the original ECG data.

Two main indexes are used to evaluate the performance of the ECG compression algorithm, i.e., CR and Percentage Root-Mean-Square Difference (PRD). CR can be calculated as (7)

$$CR = \frac{Bits\ per\ sample\ in\ the\ original\ file}{Bits\ per\ sample\ in\ the\ compressed\ file} \quad (7)$$

The PRD represents the quality of the reconstructed signal. The PRD is defined as:

$$PRD(\%) = 100 \times \sqrt{\frac{\sum_{n=0}^{N-1}\left(X_s(n) - X_r(n)\right)^2}{\sum_{n=0}^{N-1}\left(X_s(n)\right)^2}} \quad (8)$$

Where, $X_s(n)$ is the original signal and $X_r(n)$ is the reconstructed signal. CR is to evaluate the level of compression and PRD is to evaluate the information fidelity of the compression algorithm.

III. Results

Forty-eight 30-minute ECG files obtained from MIT-BIH Arrhythmia Database are used to test the performance of the algorithm. A ECG file contains ECG data of two leads, each with sampling rate of 360 samples/sec. A sampling point of the ECG signal is represented by 12 bits. The relationship between data segment length and compression effect is estimated preliminarily to obtain the optimal data segment length N. According to the format of the data packet, the compressed data length of the slow varying segments is calculated as following:

$$L_S = 5 + 12 + (N - 1) * Bit_{min} \quad (9)$$

The compressed data length of the segments with the sharp changes is calculated as:

$$L_P = 5 + 2 \times 12 + (N - 2) * Bit_{min} \quad (10)$$

It can be found that after delta encoding, the maximum absolute value of slow varying segments is generally ≤ 8, which can be expressed by a 3-bit binary number. The maximum absolute value of the data in sharp changing segments is generally not less than 64, which can be represented by a 7-bit binary number. Adding a sign bit, they can be represented by 4-bit and 8-bit binary numbers respectively. Usually, there are 1 or 2 sharp changing segments in a cardiac cycle. For the sampling rate of 360 samples/sec, the data bits of a cardiac cycle after encoding can be set as:

$$L = \mathrm{L_S} \times \left(\frac{360}{N} - 1\right) + \mathrm{L_P} \quad (11)$$

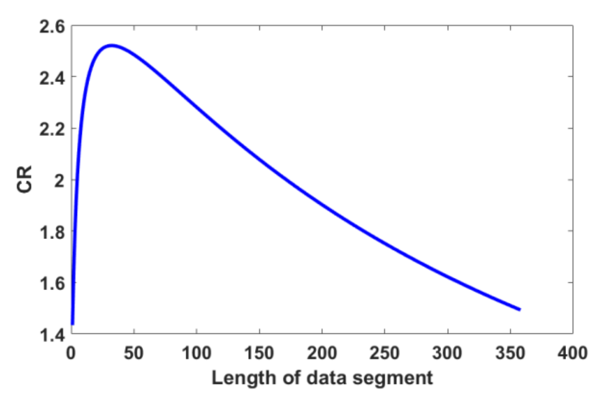

Figure 4. CR in a cardiac cycle with the length of data segment change

Figure 4 shows how the CR values vary with the data segment lengths. It can be found that the maximum CR value is 2.52 when the data segment length is located between [33:35].

The previous study has proved that the linear predictors with an order more than two do not show significant improvement on the compression [3]. Therefore, we test the first-order, the second-order and the third-order linear predictors.

$$e(n) = x(n) - x(n-1) \tag{12}$$

$$e(n) = x(n) - 2x(n-1) + x(n-2) \tag{13}$$

$$e(n) = x(n) - 3x(n-1) + 3x(n-2) - x(n-3) \tag{14}$$

As shown in Figure 5, when the order of the predictor equals to two, the CR is the largest, which indicates that the second-order linear predictor is the optimal.

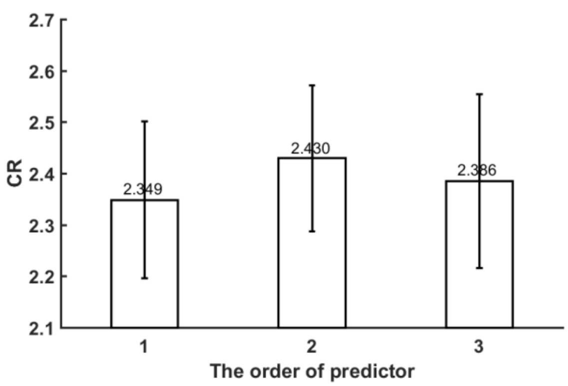

Figure 5. CR obtained by the predictor with different orders

The PRD of each data segment is 0.00%, which proves that the compression algorithm is lossless.

TABLE I. COMPARISON WITH THE EXISTING ALGORITHMS

Method	CR
Linear Predictor/Huffman Coding[11]	1.92
Linear Predictor/Dynamic Data Packaging[7]	2.28
Delta Predictor/Rice Golomb Coding[12]	2.38
This work	2.43

We also compare the performance of the algorithm in this study, with other existing low complexity algorithms that can run on low-power and low-voltage devices and results are listed in Table I. The CR of this algorithm is higher than those of the others. The results imply that the proposed algorithm in this study can increase the CR further through the combination of delta encoding and linear prediction, effective design of adaptive length encoding with optimal parameters.

CONCLUSION

In this paper, we propose a novel lossless ECG compression algorithm with low power consumption and low complexity. It includes four steps: 1) length detection, 2) delta encoding, 3) linear prediction and 4) adaptive length encoding. The algorithm is tested on forty-eight segments of 30-min ECG signals obtained from MIT-BIH Arrhythmia Database. The effect of the data segment length and the linear predictor order on CR is studied. The results show that the average CR reaches 2.43 and the PRD is as low as 0.00 when the optimal parameters are adopted, i.e., with the segment length of 33 and the predictor order of 2. It implies that the algorithm can realize lossless compression with a high compression ratio. Meanwhile, the low complexity makes it suitable for ECG monitoring of active implants.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (Grant No. 82072013, 81571761), and the Fundamental Research Funds for the Central Universities (Grant No. XJJ2015083).

REFERENCES

[1] WHO. "World health statistics 2020: monitoring health for the SDGs, sustainable development goals". World Health Organization, Geneva(2020)

[2] D Rzepka. "Low-complexity lossless multichannel ECG compression based on selective linear prediction." Biomedical Signal Processing and Control, 57(2020):101705.

[3] Tiwari, A., and T. H. Falk. "Lossless electrocardiogram signal compression: A review of existing methods." Biomedical Signal Processing and Control, 51. MAY (2019):338-346.

[4] Tomson TT, and Passman R. "The Reveal LINQ insertable cardiac monitor." Expert Rev Med Devices. 2015;12(1):7–18.

[5] Edvardsson, et al. "Use of an implantable loop recorder to increase the diagnostic yield in unexplained syncope: results from the PICTURE registry." Europace London (2011).

[6] Kannan, R., and C. Eswaran. "Lossless Compression Schemes for ECG Signals Using Neural Network Predictors." Eurasip Journal on Advances in Signal Processing, 2007.1(2007):1-20.

[7] Deepu, C. J., and Y. Lian. "A Joint QRS Detection and Data Compression Scheme for Wearable Sensors." Biomedical Engineering IEEE Transactions on, 62.1(2015):165-175.

[8] Lee, S., J. Kim, and M. Lee. "A Real-Time ECG Data Compression and Transmission Algorithm for an e-Health Device." IEEE transactions on bio-medical engineering, 58.9(2011):2448.

[9] Manikandan, M. S., and S. Dandapat. "Wavelet threshold based TDL and TDR algorithms for real-time ECG signal compression." Biomedical Signal Processing and Control, 3.1(2008):44-66.

[10] Ranjeet, et al. "Beta wavelet based ECG signal compression using lossless encoding with modified thresholding." Computers & Electrical Engineering (2013).

[11] Chua, E., and W. C. Fang. "Mixed Bio-Signal Lossless Data Compressor for Portable Brain-Heart Monitoring Systems. "IEEE Transactions on Consumer Electronics 57.1(2011):267-27

[12] Chen, S. L, et al. "Wireless Body Sensor Network with Adaptive Low-Power Design for Biometrics and Healthcare Applications." IEEE Systems Journal 3.4(2009):398-409.