A HIGH PERFORMANCE COMPRESSION ALGORITHM FOR ECG WITH IRREGULAR PERIODS

Hsiao-Hsuan Chou[1], Ying-Jui Chen[2], Yu-Chien Shiau[1,3], and Te-Son Kuo[1,4]

1 Department of Electrical Engineering, NTU, Taipei, Taiwan, R.O.C.
2 MIT, Cambridge, MA 02139, USA
3 Department of Nuclear Medicine, Far Eastern Memorial Hospital, Panchiao, Taipei, Taiwan, R.O.C.
4 Institute of Biomedicine Engineering, NTU, Taipei, Taiwan, R.O.C.

ABSTRACT

Electrocardiogram (ECG) signals have both intra- and inter-beat correlations, which can be exploited for compression by arranging the ECG signals into appropriate two-dimensional (2-D) representations. In this paper, we propose a novel approach that maps 1-D ECG signals to 2-D arrays effectively and then compresses the 2-D arrays with efficient image compression algorithms. Compared to existing 2-D ECG compression methods, the proposed algorithm is unique in that it reveals much more intra- and inter-beat correlation characteristic of ECG signals. Therefore, the image compression algorithms can achieve enhanced performance. Furthermore, unlike existing 2-D ECG compression methods, the proposed algorithm works well for both regular and irregular ECG signals with extremely varying periods. In particular, its performance is insensitive to QRS miss detection cases.

1. INTRODUCTION

Modern ECG monitoring devices generate vast amounts of data and require huge storage capacity. In order to process, transmit, and store the data efficiently, many ECG compression methods were proposed and could be classified into three major categories [1]:

1) Parameter extraction techniques: such as prediction [2] and vector quantization methods [3].
2) Transform-domain techniques: such as 2-D discrete cosine transform (DCT) [1], singular value decomposition (SVD) [4], and wavelet transforms [5][6][7].
3) Direct time-domain techniques: such as AZTEC [8], Scan-along polygonal approximation (SAPA) [9], and fan algorithm [10].

Among the categories listed above, most of the methods adopt 1-D representations for 1-D ECG signals. However, since the ECG signals have both sample-to-sample (intra-beat) and beat-to-beat (inter-beat) correlations, some 2-D compression approaches have been proposed for better compression performances. For example, Lee [1] used "cut and align beats approach and 2-D discrete cosine transform" to get pretty good compression results in regular ECG. Wei [4] used truncated singular value decomposition algorithm to compress 2-D ECG arrays. Recently, wavelets are widely used for both 1-D and 2-D ECG compression [5][6][7]. Their results are summarized in Table 1. Most of the papers showed good ECG compression performances for regular ECG cases. However, their compression performance dropped in irregular ECG signals. In order for the 2-D ECG compression algorithms to accommodate irregular ECG signals, we propose below a novel procedure that converts (1-D) ECG signals to easier-to-compress 2-D representations. As for the codec, JPEG2000 is chosen because it is the latest international standard for static image compression with better performance than others [11] and it has been adopted to compress ECG with high efficiency [12][13]. The following section will describe the 1-D to 2-D process in detail including the QRS detection and alignment, length limitation, period sorting, and mean extension steps.

2. NOVEL 2-D ECG PROCESS

Before describing the algorithm, two measures are first introduced to compare the compression performance with other algorithms: the compression ratio (CR) and the percent root mean squared difference (PRD). The CR is calculated as the number of bits in the original 1-D ECG signal over the number of bits in the compressed signal, and PRD is given by (1)

$$\text{PRD} = \sqrt{\frac{\sum_{n=1}^{N}[y(n) - \hat{y}(n)]^2}{\sum_{n=1}^{N}[y(n)]^2}} * 100\%, \quad (1)$$

where $y(n)$ is the original signal of length N, and $\hat{y}(n)$ is the reconstructed signal.

2.1. QRS detection and alignment

To map 1-D ECG signals to 2-D arrays, the peaks of QRS waves should be detected first to identify each heartbeat. Many QRS detection algorithms [14] have been proposed and we choose "A simple real-time QRS detection algorithm" [15] for its high detection accuracy (99.5%). Others with high enough detection accuracy would also work. After each QRS peak of heartbeat segments is identified, the original 1-D ECG signals are cut at every 130th sample before each QRS peak (a

130-sample shift from the QRS peak, a smooth region estimated from human physiology [16]). Note that we choose not to use the QRS peaks to delimit each heartbeat segment because we want to avoid large boundary values, which tend to result in discontinuities and make the 2-D arrays hard to compress.

2.2. Length limitation

Because the periods of heartbeats are not all the same, the lengths of the heartbeat segments can be different. This is more severe if the ECG is abnormal or the QRS detection step misses some QRS peaks such that the identified "heartbeat" is very long, resulting in a 2-D array with huge size. To deal with this issue, we need to limit the length of the detected heartbeats to form a 2-D ECG representation of reasonable sizes. We choose the length limit to be 512 pixels. Heartbeats with longer periods than the length limit will be cut into two or more segments. If the length limit is not set to 512 but 1024 or some larger value, i.e., the 2-D arrays are extended to a larger size by more dummy values such as zero or the mean value of the arrays, this will not affect the CR and PRD results because the JPEG2000 codec can process dummy values very well. The purpose of the length limit step is just to avoid unbounded array sizes if QRS cannot be detected correctly for a long time.

2.3. Period sorting

The 2-D array resulting from QRS detection and alignment already exhibits the inter-beat correlation of the original ECG signals. However, the period irregularity presents a challenge to 2-D compression algorithms. To exploit the inter-beat correlation and simultaneously obtain a superior performance, we propose an additional period sorting step which sorts the period of each heart beat segment. This is a novel and powerful method for irregular ECG compression because it reduces the period differences among the adjacent heartbeats effectively, resulting in improved CR and PRD.

Existing 2-D ECG compression methods [1][4][7] show good results for normal ECG signals but not for irregular ones with extremely varying periods. This implies that these algorithms have poorer PRD on irregular ECG signals. However, irregular signals are more significant for clinic diagnosis than normal ones. It is very desirable if a compression algorithm can process abnormal ECG signals very well. This effective period sorting step addresses this issue and better exposes the 2-D correlation structure to be exploited by the following extension method described in the next section to form a 2-D array.

2.4. Length equalization with mean extension

Several extension methods have been considered, including zero extension which pads the short segments with zeros, zero-order extension which extends a segment by repeating its last sample, and mean extensions which pads short segments with the mean of the whole signals of the last samples of heartbeat segments. In summary, this step equalizes the length of each heartbeat segment to form a proper 2-D array. Experiments indicate that mean extension with mean of the last samples of heartbeat segments performs better than the other three methods.

3. Experimental results

The proposed algorithm was applied to 100, 117, 119, and 232 in MIT-BIH arrhythmia database [17] with 11-bit resolution and 360 Hz sampling rate. Record 119 with extremely varying periods is taken for an example. First, the QRS peaks are detected, cut and aligned. Fig. 1 (a) would be the 2-D greyscale array if we were to cut the ECG at QRS peaks. Note the undesirable large values near the segment boundaries as pointed out earlier. By taking the proposed QRS detection and alignment method with zero extension (but no period sorting), we obtain Fig. 1 (b). Then, by the proposed period sorting step, Fig. 1 (c) is formed. Lastly, by mean extension with the mean of the last samples of all heartbeat segments, Fig. 1 (d) is obtained. Note that in Fig. 1, the 2-D arrays become smoother and smoother as the algorithm progresses. These 2-D representations are to be compressed by JPEG2000. The compression performances are shown in Fig. 2. It also shows that CR and PRD results are improved step by step. The comparison results of various 2-D ECG compression algorithms are summarized in Table 1. It is clear that the proposed algorithm performs better than others, especially in irregular ECG record 119. An example of the original signal, reconstructed signal, and errors are shown in Fig. 3. The errors in this case are small and evenly distributed.
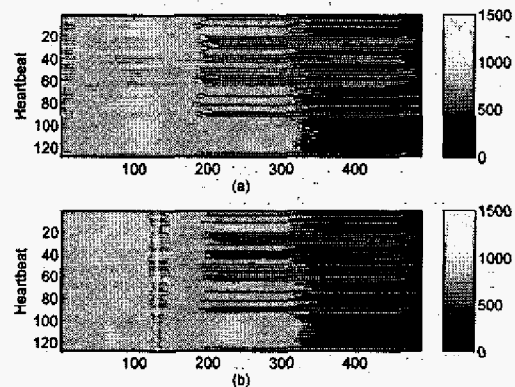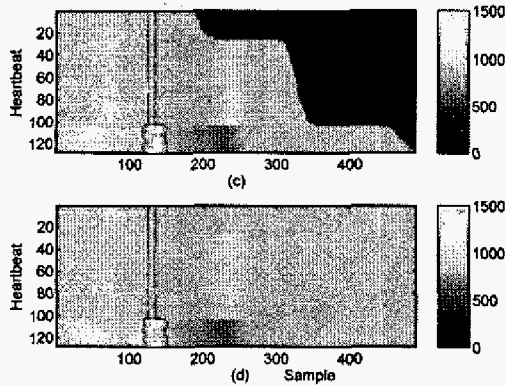
Fig. 1. Step-by-step illustrations of the 2-D greyscale arrays resulting from the proposed algorithm.
(a) Cutting at QRS peaks followed by zero extension.
(b) Cutting with 130-sample shifts from the QRS peaks followed by zero extension.
(c) Period sorting applied to (b).
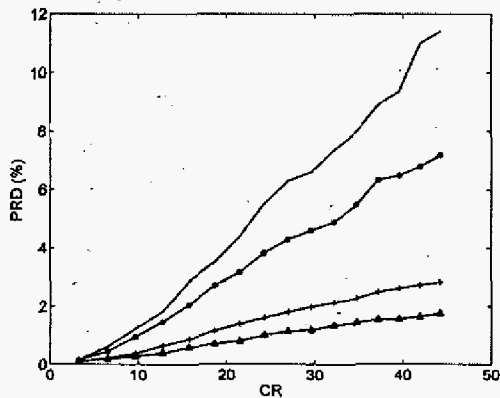(d) Period sorting applied to (b) with mean extension.

Fig. 2. The CR and PRD pairs of the various 2-D arrays.
. :   Fig. 1 (a)
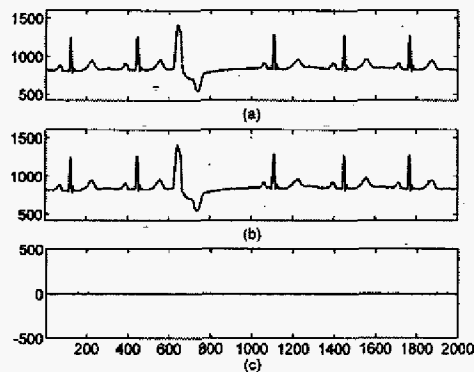*:   Fig. 1 (b)
+:   Fig. 1 (c)
△:   Fig. 1 (d)

Fig. 3. Record 119, CR=13, PRD=0.4%.
(a) Original ECG signal.
(b) Reconstructed ECG signal.
(c) Difference between (a) and (b).

4. Discussion

4.1. Overhead of original ordering and heartbeat lengths

Although the original ordering and lengths of heartbeat segments need to be recorded losslessly in the compressed data for signal reconstruction, the corresponding overhead is negligible, for example, 1/250 of the original data size of record 119 mentioned above. Thus the CR varies from 23.8 to 21.8, but PRD is improved from 5.0% to 0.81%, which justifies the slight reduction in compression ratio.

4.2. Sensitivity to accuracy of QRS detection

An interesting question to ask is what will happen when QRS peak is not detected correctly or when the irregular ECG signals do not have obvious QRS peaks in some periods. The following test case demonstrates how this approach performs with QRS miss-detections that happen frequently in abnormal ECG cases. Again, record 119 is taken for an example for comparison. The QRS detection misses 30% on purpose to make a large 2-D array without length limit (Fig. 4). The compression result is summarized in Table 1, too. Its CR and PRD pair obtained by this algorithm are 22.7 and 1.18% respectively, still much better than the other algorithms even with 30% QRS miss-detections.

Table 1. PRD results of various 2-D ECG compression algorithms.

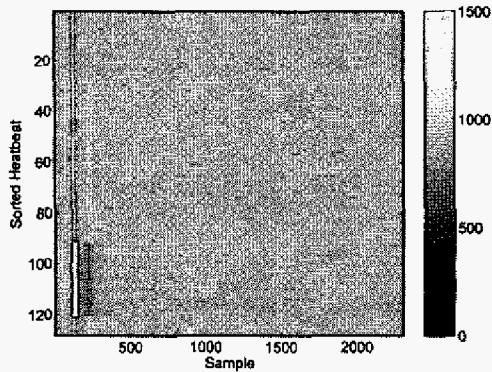| Algorithm	Record	CR	PRD (%)
A. Bilgin et. Al [13]	117	8:1	0.86
Proposed Algorithm	117	8:1	0.19
Wei et. Al [4]	117	10:1	1.18
A. Bilgin et. Al [13]	117	10:1	1.03
Proposed Algorithm	117	10:1	0.23
Lee et. Al [1]	119	24:1	10.5
Ali Bilgin et. Al [13]	119	21.6:1	3.76
Proposed Algorithm	119	21.8:1	0.81
Proposed Algorithm with 30% QRS Miss-detections	119	21.6:1	1.29

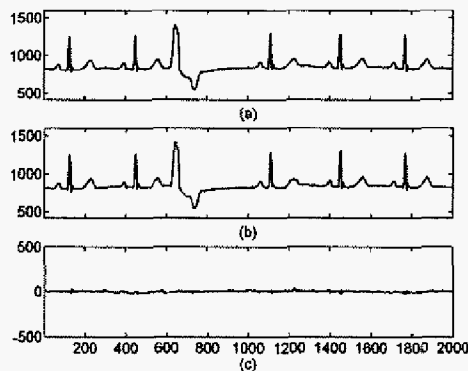Fig. 4. 2-D greyscale array of 119 with 30% QRS miss detections.

Fig. 5. Record 119 with 30% QRS miss detection (CR=22.7, PRD=1.18%).
(a) Original ECG signal.
(b) Reconstructed ECG signal.
(c) Difference between (a) and (b).

5. CONCLUSION

The proposed algorithm effectively rearranges the 1-D ECG signals to smooth 2-D images and makes it easy for JPEG2000 to enhance the compression performance. All the experimental results show how the proposed algorithm improves PRD, especially in irregular ECG and even in QRS miss-detection test cases. They also show superior performance compared to other methods in the literature.

REFERENCES

[1] H. Lee and K. M. Buckley, "ECG data compression using cut and align beats approach and 2-D transforms", *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 556-565, 1999.

[2] G. Nave and A. Cohen, "ECG compression using long-term prediction", *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 877-885, 1993.

[3] B. Wang and G. Yuan, "Compression of ECG data by vector quantization", *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 23-26, 1997.

[4] J. J. Wei, C. J. Chang, N. K. Chou, and G. J. Jan, "ECG data compression using truncated singular value decomposition", *IEEE Trans. Biomed. Eng.*, vol. 5, pp. 290-299, 2001.

[5] B. A. Rajoub, "An efficient coding algorithm for the compression of ECG signals using the wavelet transform", *IEEE Trans. Biomed. Eng.*, vol. 49, pp. 355-362, 2002.

[6] M. L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms", *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 394-402, 1997.

[7] A. R. A. Moghaddam and K. Nayebi, "A two dimensional wavelet packet approach for ECG compression", *Signal Processing and its Applications*, the 6th International, Symposium, pp. 226-229, 2001.

[8] J. R. Cox, F. M. Noile, H. A. Fozzard, and G. C. Olover, "AZTEC: pre processing program for real-time ECG rhythm analysis", *IEEE Trans. Biomed. Eng.*, vol. 15, pp. 128-129, 1968.

[9] M. Ishijima, S. B. Shin, G. H. Hostetter, and J. Sklansky, "Scan-along polygonal approximation for data compression of electrocardiograms, " *IEEE Trans. Biomed. Eng.*, vol. 30, 11, pp. 723-729, 1983.

[10] R. C. Barr, "Adaptive sampling of cardiac waveforms", *J. Electrocard.*, 21, pp. 57-60, 1988.

[11] David S. Taubman and Michael W. Marcellin, *JPEG2000: image compression fundamentals, standards, and practice*, Kluwer Academic Publishers, Boston, 2002

[12] A. Bilgin, M. W. Marcellin, and M. I. Altbach, "Compression of electrocardiogram signals using JPEG2000", *IEEE Tran. Consumer Electronics*, vol. 49, pp.833–840, Nov. 2003.

[13] A. Bilgin, M. W. Marcellin, and M. I. Altbach, "Wavelet compression of ECG signals by JPEG2000", *Conf. Data Compression*, DCC2004, pp. 527 – 527, Mar. 2004.

[14] B. U. Kohle, C. Hennig, and R. Orglmeister, "The principles of software QRS detection", *IEEE Biomed. Eng. Mag.*, pp 42-57, Jan.-Feb. 2002.

[15] J. Lee, K. Jeong, J. Yoon, and M. Lee, "A simple real-time QRS detection algorithm", *IEEE Proc. Biomed. Eng.*, vol. 4, pp. 1396–1398, 1996.

[16] J. Vander, J. H. Sherman, and D. S. Luciano, *Human Physiology*, McGraw-Holl, 6th ed, chap 14, pp 393-472, 1994.

[17] "MIT-BIH Arrhythmia Database CD-ROM", 2nd. Ed., Harvard-MIT Division of Health Sciences and Technology, Aug. 1992.