Different Approaches for Biomedical Signal (ECG) Compression in Secure Communication Network

Praveen Chandra Patel[1] Atulika[2]
[1]M.Tech Student [2]Assistant Professor
[1,2]Department of Electronics And Communication
[1,2]SRIT, Jabalpur- (M.P.), India

Abstract— Storage and transmission limitations have made biomedical signal data compression an important feature for most biomedical computerized systems. In addition, when compressed biomedical signals (ECG) or data are delivered over a public channel such as the Internet, TV etc their privacy and security would also be an important issue. Electrocardiogram (ECG) signal is a very important measure to know the Heart actual conditions so that easily found deceases. Various techniques have been proposed over the years for addressing the problem. In our paper, we are analyzing different approaches of biomedical signal compression such as ECG signal in secure communication network.ECG signals are collected both over long periods of time and at high resolution from patients. This creates substantial volumes of data for storage and transmission. Data compression seeks to reduce the number of bits of information required to store or transmit digitized ECG signals without significant loss of signal quality. A wide range of compression techniques based on different transformation techniques like MSVQ, genetic optimization, SPARSE 2D SEPARABLE TRANSFORM. DST and DCT were evaluated to find an optimal compression strategy for ECG data compression. All Testing was performed on artificially signals from the standard CSE and MIT-BIH database.

Key words: ECG, MSVQ, Sparse 2D, DST, DCT

I. INTRODUCTION

An electrocardiogram is used to monitor your heart. Each beat of your heart is triggered by an electrical impulse normally generated from special cells in the upper right chamber of your heart. An electrocardiogram — also called an ECG or EKG — records these electrical signals as they travel through your heart.

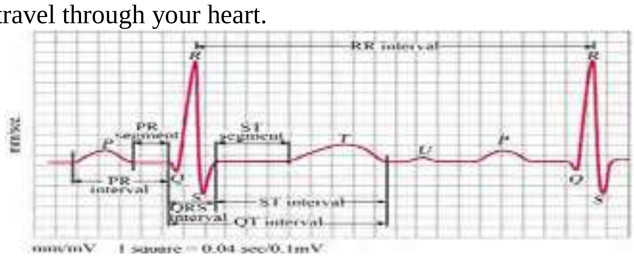

Fig.1: ECG Signal Specification

An electrocardiogram is a noninvasive, painless test. The results of your electrocardiogram will likely be reported the same day it's performed, and your doctor will discuss them with you at your next appointment.

Biomedical signals can be compressed in time domain, frequency domain, or time-frequency domain. ECG data compression algorithms have been mainly classified into three major categories [3]: 1) Direct time-domain techniques, e.g., turning point (TP), amplitude-zone-time epoch coding (AZTEC) [4], coordinate reduction time encoding system (CORTES) and Fan algorithm. 2) Transformational approaches [3], e.g., discrete cosines transformation (DCT), fast Fourier transform (FFT), discrete sine transform (DST), wavelet transform (WT) etc. 3) Parameter extraction techniques, e.g., Prediction and Vector Quantization (VQ) methods [2]. The time domain techniques which are based on direct methods were the earlier approaches to biomedical signal compression. Transform Coding (TC) is the most important frequency domain digital waveform compression method. When we compare these methods we find that direct data compression is a time domain compression algorithm which directly analyses samples where inter-beat and, intra-beat correlation is exploited. In this paper different transforms like MSVQ, Sparse 2D, DST, DCT and are studied for ECG signal compression.

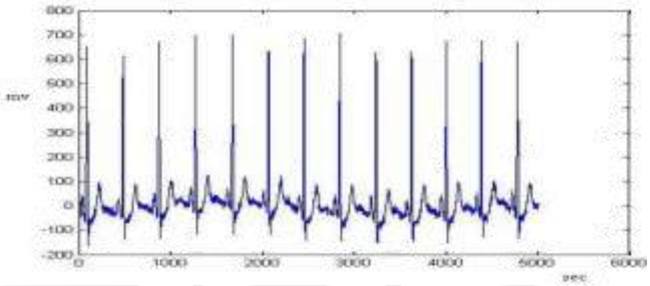

Fig.2: ECG Signal before Compression

II. EVALUATION CRITERIA & COMPRESSION RATIO

The evaluation of performance for testing ECG compression algorithms includes followings three components: First one is compression efficiency second is reconstruction error and third is computational complexity. The compression efficiency is given by compression ratio (CR).

$$\mathrm{CR} = \frac{\text{Number of samples before compression}}{\text{Number of samples after compression}}$$

III. DIFFERENT TECHNIQUE

A. MSVQ

The motivation for introducing MSVQ for ECG signal compression is determined from the observation of the waveform itself. If the time sequences of ECG signals are analyzed, it is possible to conclude that many short-length segments differ mainly in their mean values, while their wave shapes have less variability. Following this reasoning, the compression system described here defines vectors as segments from the sequence of samples in a digitized ECG signal, and then calculates and subtracts the mean value from each vector. The subsequent signal processing is as outlined above, and is summarized below in four steps.

– An ECG signal sampled at an adequate rate is appropriately segmented into vectors.

- The mean is subtracted from the vectors created in 1) and a training sequence is derived from a very large number of segments.
- From the training sequence created in 2), a VQ codebook of adequate bit rate is constructed using a standard algorithm, such as that proposed by Linde, Buzo, and Gray [39].
- Both scalar quantization of the mean value and VQ of the wave shape are included in the coding procedure foreach waveform segment. It should be noted that to decrease computational complexity, a convenient data structure for the codebook should be employed.

B. Entropy Coding

The code resulting from the MSVQ can be passed through a lossless compression method to further improve the CR. As the distribution functions of the codes can be easily obtained as a byproduct of the codebook design phase, Huffman coding can be employed. Furthermore, this post-processing stage does not affect the computational burden of the method. As an example, let us consider the case where values of R=10,K=40 and n=8 are employed. In that case an average CR of 39.78 is achieved. The use of an adaptive Huffman code will improve the results presented here at expense of increased computational load. MSVQ has been introduced for ECG signal compression, with good results in comparison with other compression methods. It does not employ any QRS detection and the bit rate of the compressed signal does not depend on sampling conditions. Practical CR's in the region of 40 have been achieved with PRD values in the range 10%–12%. This CR translates into a transmission rate in the order of 140 bit/s for real-time ambulatory monitoring. These results were obtained from an analysis that encompasses codebook size, mean word length, and vector length. An analysis of the computational loads reveals that all operations for real-time VQ are easily realizable by current DSP processors. In addition, the memory overhead due to codebook storage is of modest size, and constitutes only a small fraction of the capacity needed for storage of the bulk ECG data.

C. Sparse 2D

- The basic idea of sparse 2D is achieving an enhanced sparse representation by grouping similar segments of the input ECG signal into 2D arrays, and then using a 2D transformation (which can be complete or over complete) to transform 2D arrays. A simple justification for the effectiveness of the idea is as follow:
- Assume that the grouping is done, i.e. similar segments are placed in groups and a 1D transformation (A) is used for each group.
- In each group we have similar segments and hence after transformation we will have the same number of high magnitude coefficients for each segment in a group, say high-magnitude coefficients for each segment.
- Assuming n segments in each group, we will have n_ high-magnitude coefficients in that group. In other words this group can be represented by n_ coefficients.
- Now we should perform another 1D transform (B) on the second dimension (along each row) of each group.
- Components of this row are similar (because only similar segments are in this group), i.e. there is a kind of similarity for all members of the row.
- As an example, after using 1D transform the first or second coefficients of this row will be high magnitude (because of the compaction property of used transform). This means that the whole group can be represented by highly less than n_ coefficients instead of n_ coefficients (i.e., a much more sparse representation). Ten ECG signals with 2500 samples were used for experiments of this part. For compression we have at most 12 samples overlap. In fact based on the matching criteria defined in (2) the best matching block will be chosen between these overlapping blocks. Five methodologies are compared in this part, Wavelet db9/7, complete DCT dictionary, complete Wavelet dictionary, over complete mixed (DCT+ Wavelet) dictionary and over complete DCT dictionary. The over complete 2D DCT transform has been constructed from an over complete 2D DCT transform. The idea is based on 2D transform (complete or over complete) to enhance the scarcity of the coefficients. Our simulations show that the usage of this idea (for both complete and over complete cases) enhances the denoising results compared to soft thresholding about 4dB and extended Kalman smoother filtering about 2dB for higher input SNRs, but it does not give outstanding results for ECG signal compression.

D. Genetic Segmentation

The genetic segmentation of ECG signals discussed here can be viewed as a fundamental introductory phase for a number of detailed compression schemes. It is worth stressing that the genetic segmentation of the signal does not confine itself to a specific form of compression such as a piecewise linear approximation or any other technique by that matter. We have used the model of linear approximation realized within the individual segments. Obviously, the segmentation can serve as a starting point when proceeding with more sophisticated compression schemes, say via quadratic functions or higher order polynomials. One should emphasize that the genetic segmentation is concerned with a global optimization and this stands in sharp contrast with the most compression techniques. Simply, in genetic segmentation once we decide upon their number of segments, the optimization of the segments looks at all data *globally*. On the other hand, the existing compression techniques are *local*. No matter what specific parameters of the compression method we choose, we are not at position to determine how many segments or a compression rate we are about to develop. This quite substantially hampers our ability to use the results of compression (more specifically, the parameters of, e.g., linear approximation) to design any ECG signal classifier. For this classification, We need a feature space of a fixed dimension (that is directly implied by the number of the parameters/segments of the ECG signal that has to be the same for all signals). The existing local compression mechanisms do not guarantee the satisfaction of such classification requirement. There are several possible enhancements and extensions of the approach introduced in this study. First, which is usually a crucial point in genetic optimization; one can augment the form of the fitness function. Here the GA mechanisms were guided by the absolute values of the differences between the external values (minimal and maximal) of the derivatives or their estimates. One can envision situations (especially in case of highly noisy data) that these extreme values become too sensitive and can be highly corrupted by noise. To alleviate such shortcomings, we may use more robust estimates of the derivatives. Similarly, we can base the computations of the variability not on the extreme values of the derivatives but employ some statistical measures, say quartiles of the differences. Any combination of these measures can easily remedy the potential drawbacks of noisy data.

TABLE I
DIFFERENT ECG COMPRESSION METHODS AND THEIR REPORTED PERFORMANCE

Method	CR	PRD (%)	SF (Hz.)	ADC (bits)
TP [6]	2.0	5.3	200	12
AZTEC [7]	10.0	28.0	500	12
CORTES [11]	4.8	7.0	250	12
FAN/SAPA [8]	3.0	4.0	250	-
MSAPA/CSAPA [10]	5.0	3.5	250	8
ALZ77 [13]	13.38	-	250	8
Dual application of KLT [14]	12.0	-	250	12
Fourier descriptors [15]	7.4	7.0	250	12
Adaptive Fourier coefficient estimation [16]	16.0	3.0	500	-
Sub-band coding with extensive Markov system [19]	25.0	-	500	12
Wavelet Transform [20]	9.9	-	500	-
VQ of Wavelet coefficients [22]	10.0	5.5	360	11
Wavelet packet compression [23]	8.06	-	200	12
Peak Peaking (spline) with entropy encoding [25]	10.0	14.0	500	8

E. DCT

A discrete cosine transform (DCT) is finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications in science and engineering, from lossy compression .Where small high-frequency components can be discarded, to spectral methods for the numerical solution of partial differential equations. The use of cosine rather than sine functions is critical for compression, since it turns out that fewer cosine functions are needed to approximate a typical signal.

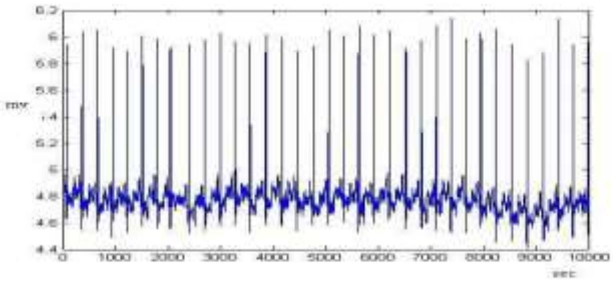

Fig.3: ECG Signal After DCT Compression

DCT-II technique, which is very advance. When there is high correlation among the input samples, which is the case in many digital waveforms including speech, music, and biomedical signals. This transform is exactly equivalent to a DFT of 4n real inputs of even symmetry where the even-indexed elements are zero.

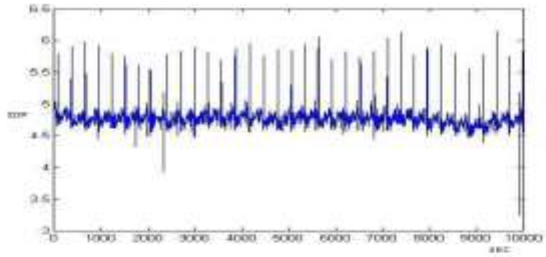

Fig. 4: ECG Signal After DCT-II Compression

F. DST

The discrete sine transform (DST) is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using a purely real matrix. It is equivalent to the imaginary parts of a DFT of roughly twice the length, operating on real data with odd symmetry (since the Fourier transform of a real and odd function is imaginary and odd), where in some variants the input and/or output data are shifted by half a sample.

A related transform is the discrete cosine transform (DCT), which is equivalent to a DFT of real and even functions. See the DCT article for a general discussion of how the boundary conditions relate the various DCT and DST types. The DST-I matrix is orthogonal (up to a scale factor). A DST-I is exactly equivalent to a DFT of a real sequence that is odd around the zero-th and middle points, scaled by 1/2. For example, a DST-I of N=3 real numbers (a,b,c) is exactly equivalent to a DFT of eight real numbers (0,a,b,c,0,−c,−b,−a) (odd symmetry), scaled by 1/2. (In contrast, DST types II-IV involve a half-sample shift in the equivalent DFT.) This is the reason for the N+1 in the denominator of the sine function: the equivalent DFT has 2(N+1) points and has 2π/2(N+1) in its sinusoid frequency, so the DST-I has π/(N+1) in its frequency.

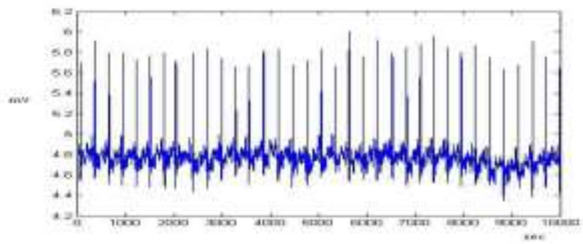

Fig.5: ECG Signal after DST Compression

Method	Compression Ratio	PRD
DCT	91.6800	0.8392
DST	70.4073	1.1967
DCT-II	94.28	1.5729

Table 2: Comparison of Dct & Dst Compression Techniques

Considering that the number of electrocardiogram records annually numbers in the millions and the use of sending electrocardiogram records over telephone lines for remote analysis is increasing, the need for effective electrocardiogram compression techniques is great. Many existing compression algorithms have shown some success in electrocardiogram compression; however, algorithms that produce better compression ratios and less loss of data

IV. Conclusions

We have studied maximum approaches for Biomedical signal (ECG) compression in secure communication network, and we found the idea is based on 2D transform (complete or over complete) to enhance the scarcity of the coefficients. By using Kalman smoother filtering about 2dB for higher input SNRs, but it does not give outstanding results for ECG signal compression. MSVQ has been introduced for ECG signal compression, with good results in comparison with other compression methods. It does not employ any QRS detection and the bit rate of the compressed signal does not depend on sampling conditions. Finally, through the use of more sophisticated VQ construction schemes and the use of entropy coding techniques, further improvements are to be expected, both in PRD-CR performance as well as in memory requirements. It is worth stressing that the genetic segmentation of the signal does not confine itself to a specific form of compression such as a piecewise linear approximation or any other technique by that matter. DST approach provides lowest CR and distortion is also high and technique of DCT improves CR with lowers PRD.

References

[1] Mean-Shape Vector Quantizer for ECG Signal Compression Juli´an L. C´ardenas-Barrera* and Juan Valent´ın Lorenzo-Ginori, *Member*, *IEEE*

[2] IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 50, NO. 10, OCTOBER 2003 1203 Communications A Genetic Segmentation of ECG Signalsucel Kocyigit, Mehmet Korurek and Bekir Karlik .

[3] S. Jalaleddine, C. Hutchens, R. Stratan, and W. A. Coberly (1990): ECG data compression techniques-*A* unified approach. *IEEE Trans. Biomed. Eng*., 37, 329-343.

[4] J. R. Cox, F. M. Nolle, H. A. Fozzard and G. C. Oliver (1968), AZTEC: A preprocessing scheme for real time ECG rhythm analysis. *IEEE Tran. Biomed. Eng.*, vol-BME-15, 128-129.

[5] M. Sabarimalai Manikandan and S. Danpat (2006): Wavelet Threshold based ECG Compression using USZZQ and Huffman Coding of DSM. In *Science Direct Biomedical Signal Processing and Control.* 261-270.

[6] B. R. S. Reddy and I. S. N. Murthy (1986): ECG data compression using Fourier descriptors, IEEE Trans. Bio-med. Eng., BME-33, 428-433.

[7] Mrs. S. O. Rajankar and Dr. S. N. Talbar (2010): An Optimized Transform for ECG Signal Compression. In *Proc. Of Int .Conf. on Advances in Computer Science*, 94-96.

[8] Shang-Gang Miaou, Heng-Lin Yen, Chih-Lung Lin (2002): Wavelet based ECG compression using Dynamic vector Quantization with Tree Code vectors in single codebook. In IEEE Transaction on Biomedical Engineering, vol. 49, no. 7, pp. 671-680.

[9] R.shanta selva Kumari, V Sadasivam (2007): A novel algorithm for wavelet based ECG signal coding. S*cience Direct Computers and Electrical Engineering*, 33, pp. 186-194.

[10] J. Abenstein and W. Tompkins (1982): A new data-reduction algorithm for real time ECG analysis. IEEE Tran. On Biomed. Engg., 29(BME-1):4, 3-8.

[11] N. Ahmed, T. Natarajan and K. R. Rao (1974): Discrete Cosine Transform. *IEEE Trans. Trans. On Computers.* C-23, 90-93.

*[12]*K. R. Rao and P. Yip (1990): Discrete cosine transform – algorithms, advantages, applications, San Diego:*Academic Press.*

[13] M. Clausen and U. Baum (1993): Fast Fourier Transforms. BI-Wiss.-Verl.

[14] L. Auslander, E. Feig and S. Winograd (1984): Abelian Semi-simple Algebras and Algorithms for the Discrete Fourier Transform. In Advances in Applied Mathematics.5, 31-55.

[15] Tinku Acharya and Ajoy K. Roy. Image Processing Principles and Applications. John Wiley.

