A NOVEL METHOD FOR ECG DATA COMPRESSION

MANJIT SINGH, BUTTA SINGH, VIJAY KUMAR BANGA, DIXT SHARMA

Abstract: Storage and transmission limitations have made electrocardiogram (ECG) data compression an important feature for most ECG computerized systems. In this paper a method for ECG data compression have been proposed. The compression algorithm has been evaluated with the MIT-BIH Arrhythmia Database. The ability of the proposed method to compress ECG signals is investigated from the results obtained by compressing and decompressing the test signals. The average percentage root mean square difference (PRD) of about 4.86%, average compression ratio (CR) of about 9 and quality score (QS) of about 1.9 was achieved on MIT-BIH data by the proposed method. The method is also compared with other compression algorithms and showed superior performance in term of PRD, CR and QS.

Keywords: ECG; Data Compression; Compression ratio; CR; PRD

Introduction: Electrocardiogram (ECG) is an important medical tool for monitoring and diagnosing a patient's heart. Further, ECG data is used to diagnose many different health disorders, including respiratory diseases, breathing disorders, and neuro-degenerative diseases such as Parkinson's and Huntington's disease. One of the largest drawbacks to the use of electrocardiography, however, is the large data size necessary to store ECG signals. Some time we need to transmit the data from one place to another using line phone or mobile for right interpretation and diagnostic .in such situations the ECG data is become mandatory to compress effectively [1]-[3]. The goal of ECG compression is to reduce the information rate as well as to preserve the relevant in the reconstructed information with low computations along with good compression ratio [4]-[7].

Compression methods can be classified in to lossy and lossless compression out of which in lossy compression smaller compressed files can be generated in comparison to lossless methods by paying the cost of some data. Further these methods can be categorised in to three types namely i) Direct methods, ii) Transformed methods and iii) Parameter extracted methods. Many algorithms have been proposed in the literature on direct data handling [8]-[13] transform method, [14]-[16] and parameter extract methods [17], [18].

Direct methds generally preserve samples that contain important information about the signal and discarding the rest. The basic idea of all these methods is to break the ECG signal into consecutive linear segments. This is done by selecting a set of significant samples in the encoder. Whereas parameter extraction methods are mainly based on linear prediction and long-term prediction methods. It is an irreversible process, where a pre-processor is employed to extract some features that are later used to reconstruct the signal.

Data : In the most emerging field of medical engineering, research subjects like cardiac arrhythmia detection, heart rate variability, ECG data compression, cardiovascular and pulmonary dynamics and artificial intelligence based medical decision support, etc. are of major interest. The Massachusetts Institute of Technology (MIT) supplies some valuable resources for such research projects. These resources include databases containing recorded physiological signals and software for analyzing, viewing and creating such recordings. The MIT-BIH sinus arrhythmia database (http://ecg.mit.edu) is an extended collection of recorded physiological signals. In the present study, 18 ECG samples from Normal Sinus Arrhythmia database are taken for inspecting the effectiveness of the proposed technique. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range.

Proposed Method: In the present work an effective protocol ECG data compression have been proposed. The proposed ECG compression technique can be divided in the compression and decompression parts. Following are the steps for compression and decompression of ECG data

- A portion of 8 ECG data samples is taken at a time from the whole ECG data samples. Now only these voltage values of ECG are sorted for further processing.
- Next step is the generation of sign bit. Starting from first ECG sample if any sample is found to be negative then one (1) and for positive, a zero (0) is taken as the symbol of sign bit for the corresponding ECG sample. The decimal equivalent to these eight binary bits is considered as the sign bit of the corresponding eight ECG data samples.
- These ECG samples are then multiplied by an appropriate large number to get integer number. Those integers are now grouped maintaining some essential logical criteria.

- The largest number from that array of 8 ECG data samples is found out and using this largest number a common amplification factor for eight samples is generated.
- Amplification factor = 9/largest number in a ECG sample array Therefore after amplification, integer part of each of the data sample will be either less than or equals to nine and a set of eight data samples can be considered as a decimal value.
- Every set of 8 ECG data samples along with the corresponding data like sign-bit and amplification factor will now be printed maintaining the following format.

Sign Bit	Decimal equivalent of Eight samples after multiplication by amplification factor	Amplification factor

In the decompression section of the proposed technique, this data is reconstructed using reverse programming logic by separating the sign-bit, amplification factor and other information.

Results and Discussion: The efficiency of proposed ECG compression technique is quantified in terms of the

- Compression ratio (CR) and
- Percentage root mean square difference (PRD).

CR is the ratio of size of original file(S_o)to size of compressed file(S_c) and is given as

$$CR = \frac{S_0}{S_c}$$

Where and are the number of bits in original ECG file and number of bits in the compressed file respectively. PRD is the measure of acceptable fidelity which is given by

$$PRD\% = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2}{\sum_{i=1}^{n} y_i^{\,2}}} \times 100\%$$

Where and represents the samples in original signal and the samples in reconstructed signal [6-7].

From the Table 1, we can see that using the proposed algorithm we can get average PRD of about 4.86%, average CR of about 9 and quality score (QS) (QS = CR/PRD) of about 1.9

Table 1. Performance of perposed method on 18 ECG data files in terms of the Compression ratio (CR), Percentage root mean square difference (PRD) and Quality score (QS)

Data	CR	PRD	QS
100	9	4.0093	2.24
101	9	4.5197	1.99
102	9	3.8841	2.317
103	9	5.2277	1.723
104	9	4.4832	2.058
105	9	5.6460	1.59
106	9	5.4012	1.66
116	9	6.2263	1.44
117	9	4.5245	1.99
118	9	5.6261	1.599
119	9	5.1137	1.759
121	9	4.5.73	1.996
122	9	4.5090	1.996
201	9	4.0790	2.201
203	9	6.1006	1.475
205	9	3.7251	2.416
208	9	5.3032	1.697
210	9	4.2999	2.083

Comparison of CR and PRD and QS of various lossy, lossless and near lossless methods with the proposed method is given in Table 2. CR of the proposed method is comparable with the methods [17, 20, 14, 21, 22. 26] and PRD of the proposed method is comparable with the methods [20, 14,23]. The CR=9 and PRD=4.86 of the proposed method is comparable with other methods but overall quality score QS=1.9 is comparable with methods [14, 20].

The Compression and decompression of the proposed method are fast and easy to implement. Simplicity of the algorithm makes one greedy to use in portable and mobile ECG data monitoring system.

Table 2. Comparison of proposed method with other methods

S. No.	Algorithm	PRD	CR	QS
1	Linear prediction interpolation and encoding [17]	3.5	7.8	2.23
2	SPHIT [19]	1.18	8	6.78
3	m-AZTEC [20]	25.5	5.6	0.22
4	Fourier descriptors 14]	7.0	7.4	1.06
5	Hilton [21]	2.6	8	3.08
6	Djohan [22]	3.9	8	2.05
7	Shrouf [23]	5.3	11.6	2.19
8	USZZQ and Huffman coding of DSM [24]	2.73	11.06	4.05
9	Fira and Goras [25]	0.61	12.74	20.89
10	Perceptual mask[26]	1.24	3.5	2.82
11	Proposed	4.86	9	1.9

Conclusion: A new method for compressing ECG data in a Lossless manner has been proposed so that after decompression there will be a very little difference between the actual and the reconstructed ECG signal. The novelty of this work is that the reconstructed signal is nearly exact reproduction of the original signal and has a moderate CR. The method achieves good CR with low distortion and morphological features are well preserved in the reconstructed signal. The proposed method performs well in term of QS also.

References:

1. I. M. Rezazadeh, S. Parvaresh, M. Erfan, H. E. Zargar, J. Proulx, “ECG Data Compression for Mobile Phone Tele-Cardiology Applications Using .NET Framework”, *1st Middle East IEEE Conference on Biomedical Engineering (MECBME),* Sharjah, 2011
2. V. Kumar, S. C. Saxena and V. K. Giri, “Direct data compression of ECG signal for telemedicine”, *International Journal of Systems Science*, vol 37, no. 1, pp 45-63, 2007
3. M. Borahan Tu¨mer, M. C. Demir, “An adaptive signal compression system with pre-specified reconstruction quality and compression rate”, *Computer Methods and Programs in Biomedicine,* vol 81, pp 99-105, 2006
4. B. Riadi, Amine, and J. Lemine, “ECG Compression Using an Ensemble Polynomial Modeling”, *Cardiovascular Engineering: An International Journal* ,vol. 4, no. 3, 2004
5. Z. Kunyan, G. Yinjing, L. Wenhong, S. Jinping, W. Xiuzhen, “A Compression algorithm for ECG based on integer lifting scheme wavelet transform”, *Journal of Electronics* ,Vol.24 , no.5, 2007
6. S.K Mukhopadhayay, S.Mitra and M.Mitra, “An ECG compression technique using ASCII character encoding”, *Measurements*, vol 45, pp 1651-1660, 2012
7. J. P. Abenstein and W.J. Tompkins, “A new data reduction algorithm for real time ECG analysis,” *IEEE Trans. Biomed. Eng*., vol.29, pp. 43-48, 1982
8. J . Sklansky and V. Gonalez, “Fast polygonal approximation of digitized curves”, *Pattern Recognisation*., vol. 12, pp. 327-331, 1980.
9. S.C. Tai, “SLOPE- A real time ECG data compressor, *Int. J. Bio- Med. Comput.*, vol. 29, pp. 175-179, 1991.
10. J. P. Abenstein and W. J. Tompkins, “New data-reduction algorithm for real-time ECG analysis”, *IEEE Trans. Biomed. Eng*., vol. BME- 29, pp. 43-48, 1982.
11. J. R. Cox, F. M. Nolle, H. A. Fozzard, and G. C. Oliver, “AZTEC, a preprocessing program for real-time ECG rhythm analysis”, *IEEE Trans. Biomed. Eng.*, vol. BME-15, pp. 128-129, 1968.
12. W. C. Mueller, “Arrhythmia detection program for an ambulatory ECG monitor”, *Biomed. Sci. Instrument*., vol. 14, pp. 81-85, 1978.

13. B. Furht and A. Perez, "An adaptive real-time ECG compression algorithm with variable threshold", *IEEE Trans. Biomed. Eng.*, vol. 35, pp. 489-494, 1988.
14. B. R. S. Reddy and I. S. N. Murthy, "ECG data compression using Fourier descriptors", *IEEE Trans. Biomed. Eng.,* vol. BME- 33, pp. 428-434, 1986
15. M. E. Womble, J. S. Halliday, S. K. Mitter, M. C. Lancaster, and J. H. Triebwasser, "Data compression for storing and transmitting ECGs/VCGs", *Proc. IEEE*, vol. 65, pp. 702-706, 1977.
16. W. S. Kuklinski, "Fast Walsh transform data-compression algorithm; ECG applications", *Med. Biol. Eng. Comput*., vol. 21, pp. 465-472, July 1983
17. U. E. Ruthann and H. V. Pipberger, "Compression of the ECG by prediction or interpolation and entropy encoding", *IEEE Trans. Biomed. Eng.,* vol. BME-26, pp. 613-623, 1979
18. G. Nave and A. Cohen, "ECG compression using long-term prediction", *IEEE Trans. Biomed. Eng*., vol. 40, pp. 877–885, 1993
19. Z. Lu. D.Y. Kim , W.A. Pearman,"Wavelet compression of ECG signals by the set partitioning in the hierarchical trees algorithm",*IEEE Trans Biomed Eng*, vol 47, no. 7.pp.849-856,2000.
20. V. Kumar, S.C. Saxena, V.K. Giri, D. Singh, "Improved modified AZTEC technique for ECG data compression: effect of length of parabolic filter on reconstructed signal", *Comput Electr Eng*, vol. 31 pp. 334–344, 2005.
21. M.L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms", *IEEE Trans Biomed Eng*, vol 44, no.15, pp.394-402,1997
22. T. Djohan, T. Q. Nguyen, and W. J. Tompkins, "ECG compression using discrete symmetric wavelet transform", *IEEE Medicine and Biology*, vol.1, pp.167-168, 1995.
23. A.A. Shrouf, M.A. Zahhad and S.M. Ahmed, "A novel compression algorithm for electrocardiogram signals based on linear prediction of wavelet coefficient", *Digital signal process.* vol. 13, pp.604-622, 2003
24. M.S. Manikdan and S. Dandapat, "Wavelet threshold based ECG compression using USZZQ and hoff man coding of DSM", *Bomedical Process Control*, vol.1, pp.261-270, 2003
25. C.M. Fira and L. Goras , " An ECG signal compression method and its valediction using NN's" ,*IEEE Trans Biomed Eng* ,vol. 55, no.4, pp.1319-1326, 2008
26. C.M.D. Rodrigo, M.M. Fabrizia and V.B. Leonardo , "Near lossless compression of ECG signal using perceptual masks in DCT domain in: *CLAIB,IFMDBE Procedding* ,vol.1, pp.229-231, 2007

* * *

Dapartment of Electronics and Communication Engineering,
Guru Nanak Dev University, Regional Campus, Jalandhar, India. manu_kml@yahoo.co.in
Amrirtsar College of Engineering and Technology,Amritsar, India.