Design and Analysis of an Efficient Technique for Compression of ECG Signal

Om Prakash Yadav, Vivek Chandra, Pushpendra Singh

Abstract- Data compression is a common requirement for most of the computerized applications. There are number of data compression algorithms, which are dedicated to compress different data formats. Even for a single data type there are number of different compression algorithms, which use different approaches. This paper examines lossless data compression algorithms and compares their performance. A set of selected algorithms are examined and implemented to evaluate the performance in compressing text data. The paper is concluded by stating which algorithm performs well for ECG Signal.

Keywords: Data compression, Lossy and Lossless Compression, ECG, Compression Ratio, Compression factor, Compression time, PRD.

I. INTRODUCTION

Compression methods have gained in importance in recent years in many medical areas like telemedicine, health monitoring, etc. All these imply storage, processing, and transmission of large quantities of data. Compression algorithms can be based on direct methods, linear transformations, and parametric methods and can be classified into two main categories: lossless and lossy.

Even though many compression algorithms have been reported so far in the literature, not so many are currently used in monitoring systems and telemedicine.

The electrocardiogram (ECG) signal is the electrical interpretation of the heart activity; it consists of a set of, well defined, successive waves denoted: P, Q, R, S, and T waves.

A great intention has been paid to the adequate and accurate analysis of the ECG signal that would lead to cardiac anomalies diagnosis [1-3].

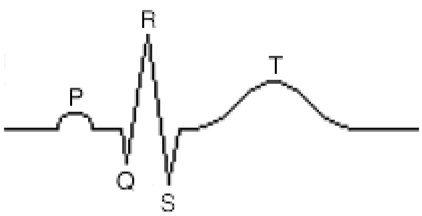

Figure1. Important feature of the ECG signal (Lynch, 1985)

Manuscript Received October 28, 2011

Om Prakash Yadav, Department of Electronics & Telecommunication Engineering, Chhatrapati Shivaji Institute of Technology, durg (CG), India. Email-opyadav@csitdurg.in.

Vivek Chandra, Department of Electrical & Electronics Engineering, Chhatrapati Shivaji Institute of Technology, durg (CG), India. Email-vivekchandra@csitdurg.in.

Pushpendra Singh, Department of Electronics & Telecommunication Engineering, Chhatrapati Shivaji Institute of Technology, durg (CG), India. Email-pushpendrasingh@csitdurg.in.

However, as the major part of real signals; the real picked-up ECG signal is corrupted by several sources of noise: EMG (electromyogram) signal (a high frequency signal related to muscle activity), the BLW (the baseline wandering: a low frequency signal caused mainly by the breathing action), the electrode motion (usually represented by a sharp variation of the baseline).This corrupted noise prevents considerably the accurate analysis of the ECG signal and useful information extraction [1].

The electrocardiogram (ECG) was introduced into clinical practice more than 100 years ago by Einthoven. It provides representation of the electrical activity of the heart over time and is probably the single-most useful indicator of cardiac function. It is widely accepted that the ECG waveforms reflect most heart parameters closely related to the mechanical pumping of the heart and can be used to infer cardiac health. The ECG waveform is recorded from the body surface using surface electrodes and an ECG monitoring system.

II COMPRESSION UTILITY

Compression techniques have been around for many years. However, there is still a continual need for the advancement of algorithms adapted for ECG data compression. The necessity of better ECG data compression methods is even greater today than just a few years ago for several reasons. The quantity of ECG records is increasing by the millions each year, and previous records cannot be deleted since one of the most important uses of ECG data is in the comparison of records obtained over a long range period of time. The ECG data compression techniques are limited to the amount of time required for compression and reconstruction, the noise embedded in the raw ECG signal, and the need for accurate reconstruction of the P, Q, R, S, and T waves. [20-22].

III COMPRESSION TECHNIQUES

Lossless compression algorithms: the Run Length Encoding Algorithm, Huffman Encoding Algorithm, Shannon Fano Algorithm, Lempel Zev Welch Algorithm, Discrete Cosine Transform, Fast Fourier Transform, Discrete Sine Transform and Discrete Cosine Transform-II are implemented and tested with a set of ECG signal. Performances of the compression methods are also evaluated at the end of the paper.

[A] Run Length Encoding

Run Length Encoding or simply RLE is the simplest of the data compression algorithms. The consecutive sequences of symbols are identified as runs and the others are identified as non runs in this algorithm.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

This Algorithm deals with some sort of redundancy [4]. It checks whether there are any repeating symbols or not, and is based on those redundancies and their lengths.

[B] Static Huffman Approach

Static Huffman Algorithms calculate the frequencies first and then generate a common tree for both the compression and decompression processes [5]. Static Huffman Encoding and Shannon Fano approaches are implemented and executed independently.

[C] Lempel Zev Welch Algorithms

The LZW method is a modification of the LZ78 approach [6]. Dictionary based compression algorithms are based on a dictionary instead of a statistical model [7-8]. This algorithm is not based on a statistical model.

[D] Discrete Cosine Transform (DCT)

A discrete cosine transform (DCT) [19] expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies [9]. Discrete Cosine Transform is a basis for many signal and image compression algorithms due to its high de-correlation and energy compaction property [10].

[E] Fast Fourier Transform (FFT)

A fast Fourier transform (FFT) [11] is an efficient algorithm to compute the discrete Fourier transform (DFT) and it's inverse [12]. An FFT is a way to compute the same result more quickly. Computing a DFT of N points in the naive way, using the definition, takes $O(N^2)$ arithmetical operations [13], while an FFT can compute the same result in only O(N log N) operations. Fast Fourier Transform is a fundamental transform in digital signal processing with applications in frequency analysis, signal processing etc [10]. The periodicity and symmetry properties of DFT are useful for compression.

[F] Discrete Sine Transform (DST)

Discrete sine transform (DST) [14] is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using a purely real matrix. DST implies different boundary conditions than the DFT or other related transforms [15].

[G] Discrete Cosine Transform-II (DCT-II)

The most common variant of discrete cosine transform is the type-II DCT [16]. The DCT-II is typically defined as a real, orthogonal (unitary), linear transformation. DCT-II can be viewed as special case of the discrete fourier transform (DFT) with real inputs of certain symmetry [17]. This viewpoint is fruitful because it means that any FFT algorithm for the DFT leads immediately to a corresponding fast algorithm for the DCT-II simply by discarding the redundant operations.

IV PERFORMANCE EVALUATION

Depending on the nature of the application there are various criteria to measure the performance of a compression algorithm [18]. Following are some measurements used to evaluate the performances of lossless algorithms.

[A] Compression Ratio (CR)

Compression ratio is the ratio between the size of the compressed file and the size of the source file [23].

$$\text{Compression Ratio} = \frac{\text{size after compression}}{\text{size before compression}} \quad (1)$$

[B] Compression factor (CF)

It is the inverse of the compression ratio. That is the ratio between the size of the source file and the size of the compressed file.

$$\text{Compression Factor} = \frac{\text{size before compression}}{\text{size after compression}} \quad (2)$$

[C] Saving Percentage (SP)

It calculates the shrinkage of the source file as a percentage:

$$\text{saving percentage} = \frac{\text{size before compression} - \text{size after compression}}{\text{size before compression}}\% \quad (3)$$

[D] Percent root mean square difference

PRD is the most prominently used distortion measure is the Percent Root mean square Difference (PRD) [18] that is given by

$$PRD = \left[\frac{\sum_{n=1}^{L_b}[x(n)-x'(n)]^2}{\sum_{n=1}^{L_b}[x(n)]^2}\right]^{(1/2)} \quad (4)$$

where x(n) is the original signal, x'(n) is the reconstructed signal and Lb is the length of the block or sequence over which PRD is calculated. PRD provides a numerical measure of the residual root mean square (rms) error.

[E] Compression Time (CT)

It is defined as the total time elapsed during the compression of original ECG signal. If the compression and decompression times of an algorithm are less or in an acceptable level it implies that the algorithm is acceptable with respective to the time factor. With the development of high speed computer accessories this factor may give very small values and those may depend on the performance of computers.

All the above methods evaluate the effectiveness of compression algorithms using file sizes. There are some other methods to evaluate the performance of compression algorithms. Compression time, computational complexity and probability distribution are also used to measure the effectiveness.

V RESULTS AND DISCUSSION

CSE database has been used to test the performance of the compression techniques. The ECG data is sampled at 333 Hz. The amount of compression is measured by CR and the distortion between the original and reconstructed signal is measured by PRD. The comparison table shown in Table 1, details the resultant compression techniques. This gives the choice to select the best suitable compression method. A data compression algorithm must represent the data with acceptable fidelity while achieving high CR.

As the PRD indicates reconstruction fidelity; the increase in its value is actually undesirable. Although DCT-II provides maximum CR, but distortion is more. So a compromise is made between CR and PRD.

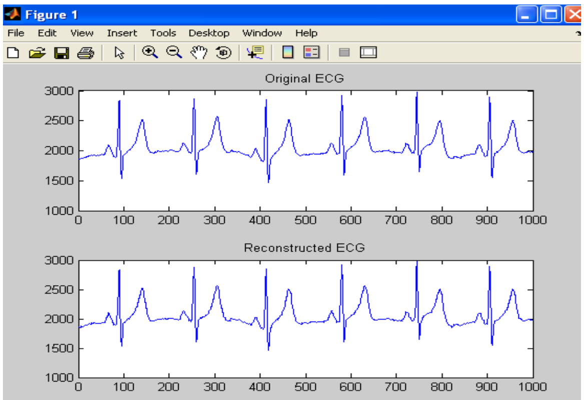

Figure 2. Fano compression analysis

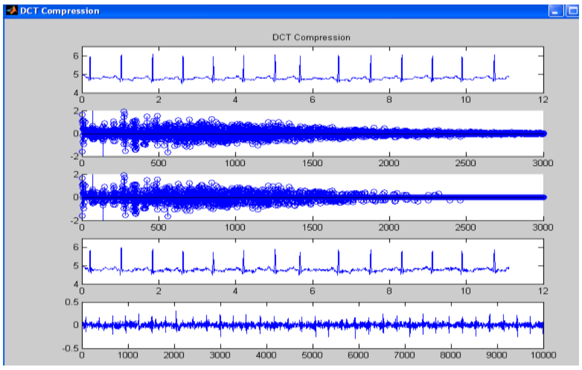

Figure 3. DCT compression analysis

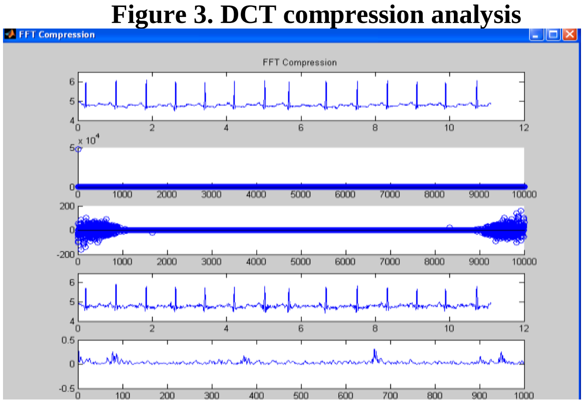

Figure 4. FFT compression analysis

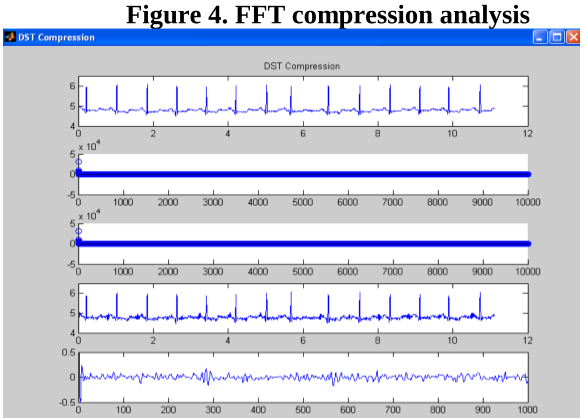

Figure 5. DST compression analysis

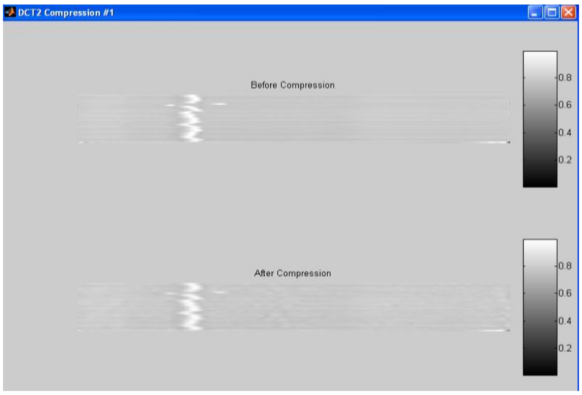

Figure 6. DCT-II compression analysis

TABLE 1. PERFORMANCE OF COMPRESSION TECHNIQUES

Method	CR	CF	SP	PRD	CT
RLE	0.384	2.60	61.60	0.616	0.123
Huffman	0.313	3.19	68.70	0.687	0.293
LZW	0.224	4.64	77.64	0.776	2.738
DCT	0.096	10.42	91.68	0.839	5.281
FFT	0.104	9.62	89.572	1.023	0.517
DST	0.148	6.76	70.407	1.196	0.517
DCT-II	0.042	23.81	94.28	1.572	0.516
Fano	0.684	1.46	31.625	0.596	0.078

VI CONCLUSION

Among the eight techniques presented, Fano provides lowest CF but distortion is low. RLE improves CF but PRD also increases. Huffman gives approximately same CF and PRD as RLE. LZW improves CF and PRD is also low. Next is FFT which gives higher CF 9.62 with PRD as 1.0237. DCT gives higher CF as 10.42 and PRD is also low as 0.839 But DCT-II provides an improvement in terms of CF of 23.81 but PRD increases up to 1.5729. Thus an improvement of a discrete cosine transform (DCT)-based method for electrocardiogram (ECG) compression is presented as DCT-II in terms of amount of compression. The appropriate use of a block based DCT-II associated to a uniform scalar dead zone quantiser and arithmetic coding show very good results, confirming that the proposed strategy exhibits competitive performances compared with the most popular compressors used for ECG compression.

REFERENCES

1. Laguna P, Jané R, Caminal P. Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE Database. Computers and Biomedical research 1994; 27:45-60.
2. Sivannarayana M. Redy DC. Biorthogonal wavelet transform for ECG parameters estimation. Medical Engineering and Physics ELSEVIER 1999; 21:167-174.
3. Bereksi-Reguig F, Chouakri SA. Computerised Cardiac Arrhythmia Detection, AUTOMEDICA. 1998;17:41-58.
4. Khalid Sayood, Introduction to data compression, Morgan Kaufmann, 3rd edition (2005).
5. Blelloch, E., 2002. Introduction to Data Compression, Computer Science Department, Carnegie Mellon University.
6. Phillips, Dwayne, "LZW Data Compression," The Computer

Applications Journal, Circuit Cellar Ink, vol. 27, June/July 1992, pp. 36-48.
7. Kesheng, W., J. Otoo and S. Arie, 2006. Optimizing bitmap indices with efficient compression, ACM Trans. Database Systems, 31: 1-38.
8. Kaufman, K. and T. Shmuel, 2005. Semi-lossless text compression, Intl. J. Foundations of Computer Sci., 16: 1167-1178.
9. N. Ahmed, T. Natarajan and K. R. Rao (1974): Discrete Cosine Transform. IEEE Trans. Trans. On Computers. C-23, 90-93.
10. 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.
11. M. Clausen and U. Baum (1993): Fast Fourier Transforms. BI-Wiss.-Verl.
12. 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.
13. Tinku Acharya and Ajoy K. Roy. Image Processing Principles and Applications. John Wiley.
14. S. Chan and K. Ho (1990): Direct Methods for computing discrete sinusoidal transforms. IEEE Proceedings, 137, 433-442.
15. G. Steidl and M. Tasche (1991): A Polynomial approach to Fast algorithms for Discrete Fourier –cosine and Fourier-sine Transforms. In Mathematics in Computation, 56 (193), 281-296.
16. E. Feig and S. Winograd (1992): Fast Algorithms for Discrete Cosine Trnsforms. IEEE Tran. On Signal Processing.vol-40(9), pp 2174-2193.
17. Xuancheng Shao and Steven G. Johnson (May 10, 2007): Type-II/III DCT/DST algorithms with reduced number of arithmetic operations. Preprint submitted to Elsevier.
18. 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.
19. K. R. Rao and P. Yip (1990): Discrete cosine transform – algorithms, advantages, applications, San Diego: Academic Press.
20. Al-Nashash, H. A. M., 1994, "ECG data compression using adaptive Fourier coefficients estimation", Med. Eng. Phys., Vol. 16, pp. 62-67
21. Bradie, Brian., 1994, "Wavelet Packet Based Compression of Single Lead ECG", Scheduled to appear in IEEE Transactions on Biomedical Engineering
22. Hamilton, Patrick S., 1991, "Compression of the Ambulatory ECG by Average Beat Subtraction and Residual Differencing", IEEE Transactions on Biomedical Engineering, Vol. 38, No. 3., pp. 253-259.
23. S. M. S. Jalaleddine, C. G. Hutchens, R. D. Strattan and W. A. Coberly. ECG Data Compression Techniques – A Unified Approach. IEEE Trans. on Biomedical Eng., vol. 37, 4 (April 1990), pp. 329-341.

AUTHORS PROFILE

Mr Om Prakash Yadav (M.Tech., B.E.) is currently working as Assistant Professor in the department of Electronics & telecommunication Engineering in Chhatrapati Shivji Institute of technology Durg, and having total 06 years experience of teaching .He Has published two research papers in national conferences and one research paper international Journals.

Dr. Vivek Kumar Chandra (Ph.D.) is currently working as HOD of Electrical and Electronics Engineering in Chhatrapati Shivji Institute of technology Durg, and having total 10 year of teaching experience. He has published 35 researsch papers in various Conferenecs(25) and journals(10).He was also awarded as **Young Scientist** in 5th Chhattisgarh Young Scientist Congress 2007 in Engineering Science by Chhattisgarh Council of Science & Technology, Raipur (CG).

Mr. Pushpendra Singh (B.E.) is working as Assistant Professor in Electronics and Telecommunication Engineering department of Chhatrapati Shivaji
ogy,durg. He is having total experience of 03 years.He
; papers in conferences.