ISSN 2229-5518

ECG Signal Compression Techniques for Data Storage and Qualitative Analysis

Nikhil Dhawan, Saurabh Gupta, Ravi Sadrani, Swati Sanghvi

***Abstract*-**ECG (electrocardiogram) is a test that measures the electrical activity of the heart. The information obtained from an electrocardiogram can be used to discover different types of heart disease. It may be useful to see how well the patient is responding to treatment. An ECG trace is a digitized version of a continuous signal. To reduce the loss in ECG signal we have used some efficient techniques in our project. Various techniques can be used for compression like the Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) etc. ECG signal being used in a wide variety of biomedical applications requires accurate results, less power requirements, faster results and low cost maintenance. Therefore compression plays a very important role in acquiring these purposes without losing the original information. In general, most of the introduced ECG compression techniques have inaccuracy and random behavior of error. Hence a new technique was proposed called as Discrete Wavelet Transform (DWT). Also from the results and computations that we have performed in our project we come to a conclusion that DWT is a better compression technique than DCT since it has better accuracy and also it correlates very well with the subjective tests.

Index terms-Accuracy, compression, Discrete Cosine Transform (DCT), Discrete Wavelet Transforms (DWT), fast, percent root mean square difference (PRD) and without losing original information

1 INTRODUCTION

Electrocardiography (ECG or EKG) is a transthoracic interpretation of the electrical activity of the heart over time captured and externally recorded by skin electrodes. It is a noninvasive recording produced by an electrocardiographic device. The etymology of the word is derived from the Greek electro, because it is related to

Electrical activity, kardio, Greek for heart, and graph, a Greek root meaning "to write". Electrical impulses in the heart originate in the sinoatrial node and travel through the intrinsic conducting system to the heart muscle. The impulses stimulate the myocardial muscle fibers to contract and thus induce systole. The electrical waves can be measured at electrodes placed at specific points on the skin. Electrodes on different sides of the heart measure the activity of different parts of the heart muscle [1],[2],[3].An ECG displays the voltage between pairs of these electrodes, and the muscle activity that they measure, from different directions, can also be understood as vectors. This display indicates the overall rhythm of the heart and weaknesses in different parts of the heart muscle. It is the best way to measure and diagnose abnormal rhythms of the heart, particularly abnormal rhythms caused by damage to the conductive tissue that carries electrical signals, or abnormal rhythms caused by electrolyte imbalances.[3] In a myocardial infarction (MI), the ECG can identify if the heart muscle has been damaged in specific areas, though not all areas of the heart are covered. The ECG cannot reliably measure the pumping ability of the heart, forwhich ultrasound-based (echocardiography) or nuclearmedicine tests are used. Compression is the reduction in size of data in order to save space or transmission time. For data transmission, compression can be performed on just the data content or on the entire transmission unit (including header data) depending on a number of factors. Compression is useful because it helps reduce the consumption of expensive resources, such as hard disk space or transmission bandwidth.Data compression methods for ECG signals have been playing an important role in computer processing and analysis of ECG. The major reason for going in for a higher compression ratio (CR) has been the desire to obtain higher density storage in medical database and hospital information systems [7].

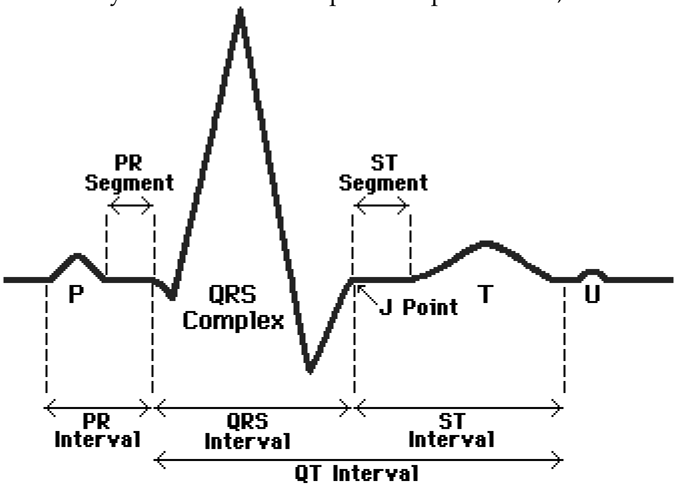
1.1Diagnosis based on heart signals

Normal sinus rhythm:

In a normal heart rhythm, the sinus node generates an electrical impulse which travels through the right and left atrial muscles producing electrical changes which is represented on the electrocardiogram (ECG) by thep-wave [1],[2],[3]. The electrical impulses then continue to travel through specialized tissue known as the atrioventricular node, which conducts electricity at a slower pace. This will create a pause (PR interval) before the ventricles are stimulated. This pause is helpful since it allows blood to be emptied into the ventricles from the atria prior to ventricular contraction to propel blood out into the body. The ventricular contraction isrepresented electrically on the ECG by the QRS complex of waves. This is followed by the T-wave which represents the electrical changes in the ventricles as they are relaxing. The

ISSN 2229-5518

cardiac cycle after a short pause repeats itself, and so on.



Therefore, on an ECG in normal sinus rhythm p-waves are followed after a brief pause by a QRS complex, then a T-wave.

Normal sinus rhythm not only indicate that the rhythm is normally generated by the sinus node and traveling in a normal fashion in the heart, but also that the heart rate, i.e. the rate at which the sinus node is generating impulses is within normal limits. There is no one normal heart rate, but this varies by age. It is normal for a newborn to have a heart rate up to 150 beats per minute, while a child of five years of age may have a heart rate of 100 beats per minute. The adult's heart rate is even slower at about 60-80 beats per minute[1],[2],[3],[6].

Sinus tachycardia:

A fast heart rate may occur with a normal heart rhythm, this is called sinus tachycardia[6]. This means that the impulse generating the heart beats are normal, but they are occurring at a faster pace than normal. This is termed sinus tachycardia and is seen normally with exercise or with excitement.

Sinus bradycardia:

The heart may slow down, yet maintain the normal pattern of rhythm (sinus), this is known as sinus bradycardia [6]. It usually is benign and may be caused by medications such as beta blockers.

Premature atrial contraction (PAC):

The atria fire an early impulse which causes the heart to beat earlier causing irregularity in the heart rhythm.

Premature ventricular contraction (PVC):

The ventricles fire an early impulse which causes the heart to beat earlier causing irregularity in the heart rhythm.

2 DIFFERENT COMPRESSION TECHNIQUES&QUALITATIVE ANALYSIS

Data compression methods for ECG signals have been playing an important role in computer processing and analysis of ECG. The major reason for going in for a higher compression ratio (CR) has been the desire to obtain higher density storage in medical database and hospital information systems. Another area wherein the need for efficient data compression for ECG has been felt is Ambulatory ECG Monitoring (AECGM). AECGM is usually done using the conventional Holter monitor which consists of a 24 hour cassette recording of the ECG. Modern Holter monitors with digital IC memory cards are now expected to improve fidelity of recording and make the system more compact [1],[2],[3],[6]. But due to the limited capacity of the IC memory cards, the sampled ECG data generated during the 24 hours has to be first compressed before it can be stored digitally. Another application which has been proposed recently is compression of ECG data for storing (along with other patient data) on a medical smart card. These and many other applications demand data compression algorithms with very high CRs.

Existing ECG data compression techniques have been classified into

a) **Direct data handling**:

The direct data handling techniques achieve data compression by removing the redundancies present in the actual ECG signal samples. Techniques such as AZTEC, CORTES, SAPA and DPCM and entropy coding come under this category.

b) **Transformation techniques**:

In contrast the transform domain techniques achieve data compression by constraining their basis functions. Many discrete orthogonal transforms such as KLT, DCT, and FT have been used for ECG data compression.

2.1 Discrete Cosine Transform

The discrete cosine transform (DCT), closely related to the DFT. The DCT's energy compaction properties are useful for applications like signal coding. The toolbox function DCT computes the unitary discrete cosine transform, or DCT, for an input vector or matrix. Mathematically, the unitary DCT of an input sequence x is

$$y(k) = w(k) \sum_{n=1}^{N} x(n) \cos \frac{\pi(2n-1)(k-1)}{2N} \quad (1)$$

$$k = 1,2 \dots N$$

Where$w(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 1 \\ \sqrt{\frac{2}{N}}, & 2 \le k \le N \end{cases}$

The DCT is closely related to the discrete Fourier transform; the DFT is actually one step in the computation of the DCT for a sequence. The DCT, however, has better energy compaction properties, with just a few of the transform coefficients representing the majority of the energy in the sequence. The energy compaction properties of the DCT make it useful in applications such as data communications [5],[7].

The function IDCTcomputes the inverse DCT for an input sequence, reconstructing a signal from a complete or partial set of DCT coefficients. The inverse discrete cosine transform is

$$x(n) = \sum_{n=1}^{N} x(n)\, y(k) \cos \frac{\pi(2n-1)(k-1)}{2N} \quad (2)$$

$$n = 1,2 \dots N$$

where

$$\omega(n) = \begin{cases} \frac{1}{\sqrt{N}}, & n = 1 \\ \sqrt{\frac{2}{N}}, & 2 \le n \le N \end{cases}$$

Because of the energy compaction mentioned above, it is possible to reconstruct a signal from only a fraction of its DCT coefficients. For example, generate a 25 Hz sinusoidal sequence, sampled at 1000 Hz.

2.2Discrete Wavelet Transform

Wavelet transform decomposes a signal into a set of basic functions. These basis functions are called wavelets [7].

Wavelets are obtained from a single prototype wavelet y(t) called mother wavelet by dilations and shifting:

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{a}} \varphi\left(\frac{t-b}{a}\right) \quad (3)$$

Where a is the scaling parameter and b is the shifting parameter

1D Wavelet Transform is given by:

$$W_f(a,b) = \int_{\infty}^{\infty} x(t) \psi_{a,b}(t) dt \quad (4)$$

Its inverse is given by:

$$x(t) = \frac{1}{C} \int_0^{\infty} \int_{-\infty}^{\infty} W_f(a,b) \psi_{a,b}(t)\, db \frac{da}{a^2} \quad (5)$$

Where$C = \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty$

2.3 QUALITATIVE ANALYSIS

ECG compression literature includes many distortion criteria for performance evaluation. Among the mostly used measures is the PRD [8],[9],[10].

$$PRD = \sqrt{\frac{\sum_{n=1}^{N} (x(n) - \tilde{x}(n))^2}{\sum_{n=1}^{N} x^2(n)}}\ \mathrm{X}\ 100 \quad (6)$$

Where

PRD is Percent Root mean square Difference

x (n) is the original signal

$\tilde{x}(n)$is the reconstructed signal and N is the length of the segment over which is PRD is calculated[11],[12],[13].In some of the articles, another version of PRD definition is used [11],[12], [13].

$$PRD = \sqrt{\frac{\sum_{n=1}^{N} (x(n) - \tilde{x}(n))^2}{\sum_{n=1}^{N} (x(n) - \bar{x}(n))^2}}\ \mathrm{X}\ 100 \quad (7)$$

Where

$\bar{x}$ is the average value of the original signal.

Equation (6) is dependent on mean value of the original segment; therefore segment with high means will exhibit an artificially low PRD.

Another distortion measure is the signal to noise ratio and is expressed as[14].

$$SNR = 10 \log_{10} \left(\frac{\sum_{n=1}^{N} (x(n) - \bar{x}(n))^2}{\sum_{n=1}^{N} (x(n) - \tilde{x}(n))^2} \right) \quad (8)$$

Relation between SNR & the PRD (7) is

$$SNR = 40 - 20 \log_{10}(PRD) dB$$

Or

$$PRD = 10^{\left(\frac{-SNR}{20}\right)}\ \mathrm{X}\ 100 \quad (9)$$

Cross Correlation (CC) measures that determine the amount of similarity between original & reconstructed signal is another objective criterion for testing the quality of reconstructed signal. It is given by

$$CC = \frac{\sum_{n=1}^{N} (x(n) - \bar{x})(\tilde{x}(n) - \bar{\tilde{x}})}{\sqrt{\sum_{n=1}^{N} (x(n) - \bar{x})^2} \sqrt{\sum_{n=1}^{N} (\tilde{x}(n) - \bar{x})^2}} \quad (10)$$

3 RESULTS

Different ECG signals from MIT/BIH arrhythmia database were analyzed using MATLAB software. And the various parameters thatwere used for analysis are PRD, SNR & CC.

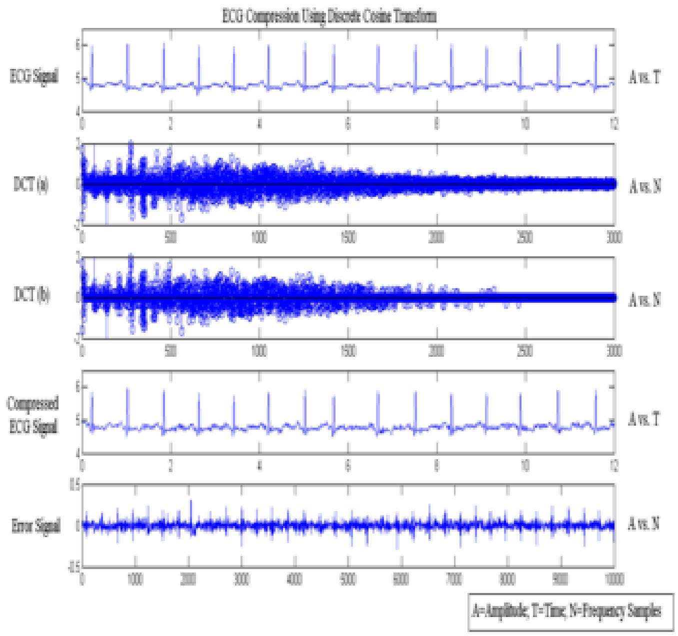


Fig. 1 Simulated result of DCT Compression technique1(a):Normal ECG Lead II Wave
1(b,c): DCT of ECG Wave by different sampling periods
1(d): IDCT and last one is the error signal.

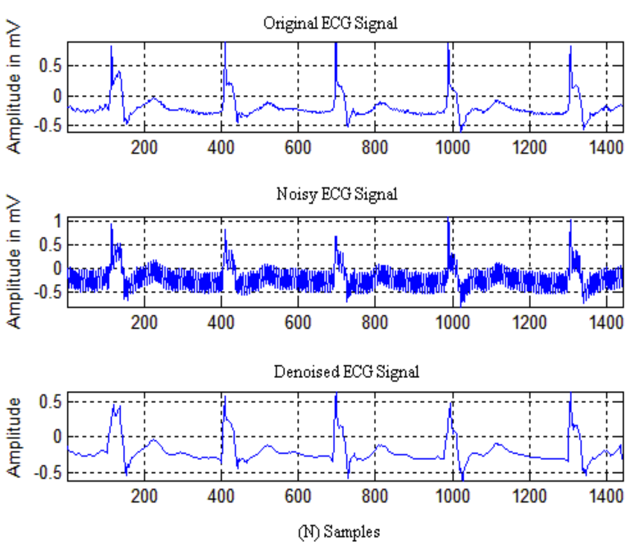


Fig 2 Simulated result of DWT Compression technique
2(a): Original ECG signal
2(b): Contaminated signal while transmission
2(c): Compressed ECG signal

4 CONCLUSION AND OBSERVATION

Electro cardiogram (ECG) is used for the measurement of electrical activity of the heart. In our report we study the important parameters of ECG signal.

ECG signal being used in a wide variety of biomedical applications requires accurate results, less power requirements, faster results and low maintenance cost.

Therefore compression plays a very important role in acquiring these purposes without losing the original information. To evaluate the performance of these algorithms, several objective and subjective criteria are available in ECG literature. Performance evaluation of the considered measure shows that it almost matches the results obtained using subjective evaluation tests. Results indicate that by using the above techniques the use of online data communication schemes can be enabled with high compression without sacrificing the quality of the transmitted signals.

According to our theoretical analysis, and according to what study we have done on the compression techniques, we conclude that DWT compression is highly efficient than DCT compression. DWT gives a signal with the most accurate results after compression and still gives a far better compression ration than DCT.

Most of the biomedical data compression methods have been developed for ECG signals. However, these methods can be applied to other biomedical signals with some modifications. It is difficult to compare the efficiency of biomedical data compression schemes because coding results of various data compression schemes are obtained under different recording conditions such as sampling frequency, bandwidth, sample precision, and noise level, which may drastically affect the currently used performance measures.

According to our evaluation compression DCT Ratio is 90.430.

5 ACKNOWLEDGMENT

The authors express sense of gratitude to respected Mr. Hemant Kasturiwale, Assistant Professor, Department of Electronics Engineering, Thakur College of Engineering & Technology , for his inspiring guidance and being patient enough to make us understand our mistakes and being tenacious enough all the time to rectify them. His support was admirable all the time till we could get every bit of it. It would have been impossible without his support and guidance all throughout.

6 REFERENCES

ISSN 2229-5518

[1] "A Comparison of Single Lead ECG Data Compression Techniques"http://www.hcs.harvard.edu/~weber/HomePage/Papers/ECGCompression.

[2]LaiaBayarri Portol'es2009-09-15LiTH-IMT-ERASMUS-R-09/34-SE , "Lossless Compression of ECG signals" Performance Analysis in a Wireless Network.

[3] Victor-Emil Neagoe, Iuliana-FlorentinaIatan and SorinGrunwald, "A Neuro-Fuzzy Approach to Classification of ECG Signals for Ischemic Heart Disease Diagnosis"

[4] Chia-Hung Lin, "Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier" Computers & Mathematics with Applications archive Vol.55, Issue4, February2008.

[5] J.A. Lay, Ling Guan, "Image retrieval based on energy histograms of the low frequency DCT coefficients,"icassp, vol.6,pp.3009-3012, Acoustics, Speech, and Signal Processing, 1999.Proceedings vol.6,1999 IEEE International Conference on, 1999.

[6] Abstracts of NHLBI-Supported Research, 1972 - Present (links to the NIH CRISP Database

[7] Data Compression:The Complete Reference David Solomon – Springer (4th Edition – Dec 2006)

[8] S.G.Miaou, H.L.Yen, and Ch.L.Lin, "Wavelet based ECG compression using dynamic vector quantization with tree code vectors in single codebook," *IEEE Trans.Biomedic.Eng.*.vol.49,no.7,pp.671-680,Jul.2002

[9] M.L.Hilton, "wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans.Biomedic.Eng.*.vol.44,no.5,pp.394-402,May1997.

[10] Z.Lu, D.Y.Kim and W.a.Pearlman, "Wavelet compression of ECG signals by the set portioning in hierarchical trees(SPIHT)algorithm," *IEEE Trans.Biomedic.Eng.*.vol.49,no.7,pp.849-856,Jul.2002

[11] A.Alshamali and A.S. Al-Fahoun, "Comments on 'an efficient coding algorithm for the compression of ECG signals using the wavelet transform," *IEEE Trans.Biomedic.Eng.*.vol.50,no.8,pp.1034-1037,Aug.2003.

[12] Y. Zigel, A.Cohen, and A.Katz, "ECG signal compression using analysis by synthesis," *IEEE Trans.Biomedic.Eng.*.vol.47,no.10,pp.1308-1316,Oct.2000.

[13] "The weighted diagnostic distortion(WDD)measure for ECG signal compression," *IEEE Trans.Biomedic.Eng.*.vol.47,no.11,pp1422-1430,Nov.2000.

[14] M. Ishijima, "Fundamentals of the decision of optimum factors in the ECG data compression," *IEEE Trans.Inf.Sys.*.vol.E76-D,no.12,Dec.1993.

- *Nikhil Dhawan is currently pursuing degree program in electronic engineering in Thakur College of Engineering and Technology from University of Mumbai, India, PH- +919029495025.*
  E-mail: nikhil.h.dhawan@gmail.com
- *Saurabh Gupta is currently pursuing degree program in electronic engineering in Thakur College of Engineering and Technology from University of Mumbai, India, PH- +919820449145.*
  E-mail: 100rabhgupta11@gmail.com
- *Ravi Sadrani is currently pursuing degree program in electronic engineering in Thakur College of Engineering and Technology from University of Mumbai, India, PH- +919870073052.*
  E-mail: ravi.sd101@gmail.com
- *Swati Sanghvi is currently pursuing degree program in electronic engineering in Thakur College of Engineering and Technology from University of Mumbai, India, PH- +919773240450.*
  E-mail:swati13sanghvi@gmail.com