

# A Review on Various Types of ECG Data Compression Techniques

Himanshu Kumar<sup>1</sup>, Kamal Niwaria<sup>2</sup>, Dr Bharti Chourasia<sup>3</sup>

<sup>1</sup>M Tech Scholar Dept. of Electronics and Communication RKDF IST SRK UNIVERSITY, BHOPAL

<sup>2</sup>Asst. Prof. Dept. of Electronics and Communication RKDF IST SRK UNIVERSITY, BHOPAL

<sup>3</sup>HODS Dept. of Electronics and Communication RKDF IST SRK UNIVERSITY, BHOPAL

\*\*\*

## ABSTRACT:

Electrocardiogram (ECG) data compression reduced the storage requirements to develop a more efficient tele cardiology system for cardiac analysis and diagnosis. The ECG compression without loss of diagnostic information is based on the fact that consecutive samples of the digitized ECG carry redundant information that can be removed with very less computing effort. This paper focuses on providing a comparison of the major techniques (direct, transform, parameter extraction and 2D approaches) of ECG data compression which are intended to attain a lossless compressed data with relatively high compression ratio (CR) and low percent root mean square difference (PRD). The paper concludes with the presentation of a framework for evaluation and comparison of ECG compression schemes..

## Keywords

Electrocardiogram; ECG; compression; CR; PRD; PRDN; QS

## 1. INTRODUCTION

The non-invasive extraction of physiological and clinical information hidden in biomedical signals is an important and fascinating field of research. Noninvasive assessment of the physiological parameters of a patient enables to study the physiology and patho-physiology of the investigated system, with minimal interference and inconvenience. Endogenous biomedical signals from physiological systems are acquired for a number of reasons including diagnosis, post surgical intensive care monitoring, neonatal monitoring and guide therapy and for research. The electrocardiogram (ECG) is a non-stationary signal containing information about the physiological condition of the heart. The electrical activity of the heart depicts the morphology and durations of the P-QRS-T intervals (Figure 1). The P, QRS complex and T features of ECG reveal the rhythmic depolarization and re polarization of the myocardium contractions of heart's atria and ventricles [1]. The time intervals between various peaks contain clinical information about the nature of possible disease afflicting a heart [2].

Due to low cost and non-invasion, ECG signal has been extended for heart disease diagnosis and ambulatory monitoring resulting in enormous volume of the data. In course of a 24-h ECG observation or multichannel biological signal acquisition, real-time data compression methods are required for the effective use of communications channels such as wired channel, wireless environment and cloud computing. The ECG data compression is also required for the transmission of ECG signals across intensive care units, emergency tele-medical services, telemedicine, home care, space programs, sports, military, public telephone networks, cellular networks and wireless communication systems [4-5]. ECG is having possibility of redundant information reduction

In general, ECG compression can be classified into lossy and lossless techniques [7]. The lossless compression guarantee the integrity of reconstructed data while compromised compression ratio (CR), with nearly 0% reconstruction error, on the other hand lossy compression is having high CR with varying level of reconstruction error [6].

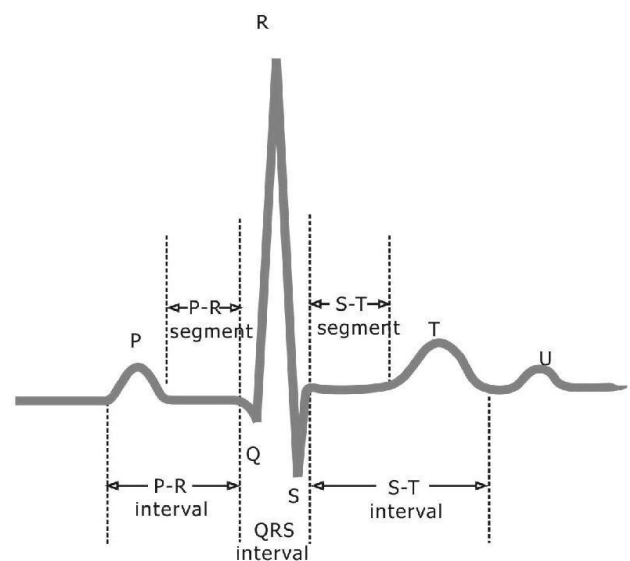


Fig 1: Time intervals of ECG

ECG signal compression techniques widely fall into three categories of direct method, transformation method and parameter extraction method [7, 8]. The direct data compression method openly analyzes and reduces data points in the time domain and the example includes turning point (TP) [25], amplitude zone time epoch coding (AZTEC) [3], Improved modified AZTEC technique [9], coordinate reduction time encoding system (CORTES), SLOPE [10], the delta algorithm and the Fan algorithm [11]. The transformed method analyzes energy distribution by converting the time domain to some other domain and example includes Fourier transform, Fourier descriptor [12], the discrete cosine transform (DCT) [13], DCT with modified stages [14, 15] and wavelet transform [16], and the compressed sensing [17]. The parameter extraction method is based upon dominant feature extraction from raw signal; examples include neural based or syntactic methods [18], peak picking and linear prediction method [19]. The other methods for compression includes ASCII based encoding for

Incorporation of ECG data as ASCII character in existing technology [20-23].

This paper explores various methods and approaches proposed earlier in literature for ECG compression and provide performance comparison of these techniques in terms of evaluation parameters.

## 2. PERFORMANCE EVALUATION

The validation of an ECG compression technique can be judged by compression efficiency and error Criterion. These parameters measure the ability of compression technique to reconstruct the signal and to preserve the relevant information. To verify reconstruction signal quality following performance parameters are used [9, 13, 20].

### 2.1 Compression Ratio (CR)

It is defined as the ratio of the original signal size and compressed signal size. The CR provides information about the degree by which the compression algorithm removes the redundant data. Higher the CR less number of bits required to store or transmits the data which can be defined as

$$CR = \frac{B_o}{B_c}$$

Where,  $B_o$  is the total number of bits required to represent original data and  $B_c$  total number of bits required to represent compressed data.

### 2.2 Percent Mean Square Difference (PRD)

It measures the error between original and reconstructed signal.

$$PRD(\%) = 100 \sqrt{\frac{\sum_{n=1}^N (X_s(n) - X_r(n))^2}{\sum_{n=1}^N (X_s(n))^2}}$$

Where  $N$  is the no of data samples,  $X_s(n)$  is the original and  $X_r(n)$  is the reconstructed signal.

### 2.3 Percentage Root Mean Square Difference Normalized (PRDN)

It is normalized version of PRD, which is independent of the signal mean value  $\bar{X}$ .

$$PRDN(\%) = 100 \sqrt{\frac{\sum_{n=1}^N (X_s(n) - X_r(n))^2}{\sum_{n=1}^N (X_s(n) - \bar{X})^2}}$$

### 2.4 Quality Score (QS)

It is the ratio of CR and PRD which quantifies the overall performance of the compression technique.

$$QS = \frac{CR}{PRD}$$

### 2.5 Root Mean Square Error (RMS)

It provides measure of error in reconstructed signal with respect to original signal.

$$RMS(\%) = 100 \sqrt{\frac{\sum_{n=1}^N (X_s(n) - X_r(n))^2}{(N-1)}}$$

## 2.6 Signal to Noise Ratio (SNR)

It is the measure of degree of noise energy introduced by compression in decibel (dB) scale

$$SNR = 10 \log \left( \frac{\sum_{n=1}^N (X(n) - \bar{X})^2}{\sum_{n=1}^N (X_s(n) - X_r(n))^2} \right)$$

## 3. ECG DATA COMPRESSION TECHNIQUES

ECG compression methods are classified as: lossless and lossy. In lossless method, compressed signal is reconstructed in exact form of original signal and in lossy method; compressed signal is reconstructed with some error. Another classification is based on techniques applied for compression and can be categorized as:

### 3.1 Direct Time-Domain Techniques

Direct methods are based on the extraction of a subset of significant samples. Direct time-domain ECG compression techniques have efficient performance in terms of processing speed and CR. These techniques explore the redundancies present directly in the ECG samples. Direct compression techniques can be based on three approaches: tolerance comparison compression, differential pulse code modulation (DPCM), and entropy coding [3]. Next, the significant work that has been focused towards direct ECG compression techniques is discussed.

#### 3.1.1 AZTEC technique

coxetal. Developed the AZTEC algorithm for preprocessing of real-time ECG's for rhythm analysis [3]. It has become a popular data reduction algorithm for ECG monitors. The AZTEC algorithm converts raw ECG sample points into plateaus. The amplitude value and length of each plateau are stored for reconstruction. Although the AZTEC technique is capable to compress with CR of 10 but the reconstruction error is not clinically acceptable. The step-like reconstructed signal may misinterpret the ECG features especially in the slow varying slopes of P and T peaks of the ECG [3, 24].

#### 3.1.2 TP algorithm

The TP is direct data compression technique to reduce the ECG sampling frequency without diminishing the elevation of large amplitude QRS complexes [25]. The TP algorithm achieved fixed CR of 2, with almost zero reconstruction error *i.e.* reconstructed signal resemble the original ECG signal. The drawback of this technique is its unsuitability for equally spaced time intervals.

#### 3.1.3 CORTES technique

CORTES algorithm is a hybrid approach of AZTEC and TP to achieve high CR of the AZTEC and the low reconstruction Error of the TP technique. TP was applied to the high frequency QRS region and AZTEC to the iso-electric regions of the ECG. The performance analysis of the AZTEC, TP, and CORTES based compression for ECG's at 200 Hz sampling frequency produced CR of 5, 2 & 4.8 and PRD's of 28, 5 & 7 respectively.

### 3.1.4 Fan and SAPA technique

Garden hire proposed and evaluated the Fan method for ECG compression [11]. Both Fan and Scan-Along Polygonal Approximation (SAPA) techniques are the first-order interpolation with two degrees of freedom (FOI-2DF) algorithms for ECG compression [24]. The Fan method implements the FOI-2DF without storing all the actual data points between the last transmitted and the present point.

### 3.1.5 Improved Modified AZTEC

In modified AZTEC algorithm, adaptive statistical parameters of the signal to be compressed are calculated. The adaptive algorithm optimizes the tradeoff between CR and PRD. [9]. The CR ranges between 2.76 and 9.91 for the threshold variation from 0.010 to 0.035 and PRD from 4.54 to 7.99 was achieved.

### 3.1.6 Delta coding

A technique called delta coding with threshold was proposed for three-lead ECG compression. ECG samples of three-lead ECG signals with successive differences more than a threshold value are stored. Otherwise samples are considered redundant and removed. In other words, actual ECG samples are replaced by the first-difference signal (amplitude between successive samples). The DPCM based ECG compression system consists a quantizer in the compression stage and an estimator in both the compression and the reconstruction stages.

## 3.2 Transform-Domain Techniques

Transform based ECG compression techniques are performed by the application of linear orthogonal transformation to ECG samples. Thus original samples of ECG are subjected to a transformation and the compression is performed in the entirely new domain like Fourier transform (FT), DcT and wavelet etc [8, 12, 16]. These techniques pose higher CR than direct techniques and are insensitive to noise present in ECG signals.

### 3.2.1 FT domain

With FT the frequency-amplitude representation of the signal is obtained [8, 12]. To reconstruct the signal inverse FFT is applied. Limitation of FT is that it fails to provide the information regarding the exact location of frequency component in time.

### 3.2.2 DCT domain

DCT represents a signal as a sum of varying magnitude and frequency. It implies different boundary condition & often used in signal and image processing for lossy data compression. DCT has strong "energy compaction property" & provide high de-correlation. In DCT compression signal information can restore in a restrict number of DCT coefficients [15]. DCT-II provide very impressive CR but at the cost of high distortion.

### 3.2.3 Wavelet domain

Wavelet transform have the beauty to analyze the signal both in time as well as frequency domains simultaneously. Many researchers concentrate on wavelet based ECG compression techniques. Recently, one-dimensional (1D) and two-dimensional (2D) Wavelet transform based ECG compression techniques with impressive CR, low PRD and smooth signal quality are presented in literature. The 2D compression approaches achieve high CR but involve complex steps; accurate QRS detection, period normalization, amplitude normalization and mean removal etc.

### 3.2.4 JPEG 2000

JPEG2000 is the latest image compression standard applied to the compress ECG signals. Some compression techniques use JPEG2000 as the encoder for the 2-D ECG image with high CR.

## 3.3 Parameter Extraction Techniques

These are irreversible processes which retain the particular characteristics or parameters of the ECG signals. The parameter extraction method is based upon dominant feature extraction from raw ECG signal; examples include neural based or syntactic methods [18], peak picking and linear prediction method.

### 3.3.1 Peak picking approach

The ECG signal is sampled at peaks and significant fiducially points are extracted. Imai et al. presented an ECG peak-picking compression system by reconstructing the signal using spline functions

### 3.3.2 Long term prediction approach

Long-Term Prediction (LTP) explores the "periodicity" property of ECG signal to reduce redundancy and to achieve high CR. Reconstruction error; PRD for LTP approach is lower than the conventional linear short term prediction (STP)

## 3.4 2D Compression Techniques

In most of the conventional compression techniques listed above, 1D representation of ECG signals has been adopted. These techniques explore the sample-to-sample (intra-beat) correlations properties for ECG compression. However, since the ECG signals have both intra-beat as well as beat-to-beat (inter-beat) correlations, 2D representation of ECG may produce better compression performance.

Many researchers have proposed ECG compression techniques by treating 1D ECG signal as a 2D image and exploiting the inter- and intra-beat correlations by encoder. These techniques broadly consists steps: QRS detection, ECG segmentation, preprocessing, and transformation. The "cut and align beats approach and 2D DCT" and "period normalization and truncated SVD algorithm" are available preprocessing techniques to get good compression results in ECG. This kind of preprocessing is also often associated with the use of state- of-the-art image encoders, like JPEG2000. A lossy compression technique based on converting the 1D ECG signal into 2D ECG image have been proposed. A period sorting preprocessing technique was introduced, which consists of a length-based ordering of all periods. The authors exploited inter and intra-beat dependencies to compress irregular ECG signals. The technique is based on the supposition that periods with similar lengths tend to be highly correlated, which is not a very strong assumption and may not be valid for pathological ECG signals. Another preprocessing technique consists of QRS detector, period length normalization, period preprocessing and image transform was proposed in. The technique focused on the preprocessing stage by reducing the vertical high-frequency content of the 2D image. The image make smoother by reducing sharp discontinuities along the vertical direction with variance based complexity sorting approach. The authors rearranged the periods in such a way that make the image smoother by putting the minimum variant period on the top of the image then compare other periods with this selected period using

Mean squared error, and arrange the most similar periods near to the top.

**Table 1 Performance comparison of different ECG compression techniques**

Compression Technique	CR	PRD	QS
AZTEC [3]	10	28	0.3571
Improved modified AZTEC [9]	2.76-9.91	4.54-7.99	0.6079-1.24
FAN/SAPA [11]	3.0	4.0	0.75
SLOPE[10]	4.8	7.0	0.685
TP [25]	2	5.1	0.392
SPIHT [16]	8	1.18	6.779
Al-Shrouf [19]	11.6	5.3	2.188
Perceptual Masks [15]	3.5	1.24	2.822
Quantized DCT coefficients [13] (Min CR) (Max CR)	6.2 10.9	1.5 3	4.13 3.63
Mukhopadhy [21]	15.72	7.89	1.992
Mukhopadhy [20]	7.18	0.023	312.173

#### 4. PERFORMANCE COMPARISON

The main objective of any compression technique is to attain maximum CR while preserving the significant signal morphology features with minimum PRD. To achieve better performance various compression techniques are proposed in literature and discussed in previous section. This section provides the comparative analysis of efficiency and performance of these techniques in terms of CR and PRD as shown in Table 1.

#### CONCLUSION AND FUTURE SCOPE

This paper gives an overview of ECG compression techniques, with a bibliographical survey of relevant background, applications, limitations and the present state. It is based on many research articles published from the past decades. 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 most important reason seems to be the fear that the recovery distortions produced by compression methods with loss of information might lead to erroneous interpretations.

Future directions for research may include the incorporation of joint applications of direct, transform and parameter extraction approaches that will allow us to further reduce the PRD (lossless compression) with improved CR and QS.

#### REFERENCES

- [1] Houghton, A. R. and Gray D. 2003. Making sense of the ECG: A Hands-on Guide. Aold Publishing company.
- [2] Singh B., Singh D., Jaryal A.K. and Deepak K.K. 2012. Ectopic beats in approximate entropy and sample entropy-based HRV assessment. *International Journal of Systems Science*, 43(5), 884-893.
- [3] Cox J.R., Nolle F.M., Fozzard H.A., and Oliver G.c. 1968. AZTEc, a preprocessing program for real-time ECG rhythm analysis. *IEEE Trans. Biomed. Eng.*, 15, 128-129.
- [4] Mitra M., Bera J.N. and Gupta R. 2012. Electrocardiogram compression technique for global system of mobile-based offline telecardiology application for rural clinics in India. *IET Sci. Meas. Technol.*, 6(6), 412-419.
- [5] Kulkarni P.K., Kumar V. and Verma H.K. 1997. Direct data compression techniques for ECG signals: effect of sampling frequency on performance. *International Journal of Systems Science*, 28(3), 217-228.
- [6] Koski A. 1997. Lossless ECG encoding. *comput. Methods and Programs Biomed.*, 52(1), 23-33.
- [7] Kumar V., Saxena S.c. and Giri V.K. 2006. Direct data compression of ECG signal for telemedicine. *Int. J. Syst. Sci.*, 37(1), 45-63.
- [8] Shinde A. and Kanjalkar P. 2011. The comparison of different transform based methods for ECG data compression. *Proceedings of International conference on Signal Processing, communication, computing and Networking Technologies*, 332-335.
- [9] Kumar V., Saxena S.c., Giri V.K. and Singh D. 2005. Improved modified AZTEc technique for ECG data compression: Effect of length of parabolic filter on reconstructed signal. *comput. Electr. Eng.*, 31,334-344.
- [10] Tai S.c. 1991. SLOPE- A real-time ECG data compressor. *Int. J. Bio-Med. comput.*, 29(2), 175-179.
- [11] Barr R.c., Blanchard S.M., and Dipersio D.A. 1985. SAPA-2 Is the Fan. *IEEE Trans. Biomed. Eng.*, 32, 337.
- [12] Reddy B.R.S. and Murthy I.S.N. 1986. ECG data compression using Fourier desCRiptors. *IEEE Trans. Biomed. Eng.*, 33(4), 428-434.
- [13] Batista L.V., Melcher E.U. and carvalho L.c. 2001. compression of ECG signals by optimized quantization of discrete cosine transform coefficients. *Med. Eng. Phys.*, 23(2), 127-34.
- [14] Lee S., Kim J. and Lee Jong-Ho. 2011. A real-time ECG data compression and transmission algorithm for an e-health device. *IEEE Trans. Biomed. Eng.*, 58(9), 2448-2455.
- [15] Duarte R.c.M., Matos F.M. and Batista L.V. 2007. Near-lossless compression of ECG signals using perceptual masks in the DcT domain, *IFMBE Proceedings*, 18, 229-231.
- [16] Lu Z., Kim D.Y., and Pearlman W.A. 2000. Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm(SPIHT). *IEEE Trans. Biomed. Eng. Wavelet*, 47(7), 849-856.

- [17] Mamaghanian H., Khaled N., Atienza D., and Vanderghenst P. 2011. compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes. *IEEE Trans. Biomed. Eng.*, 58(9), 2456–2466.
- [18] Iwata A., Nagasaka Y., and Suzumura N. 1990. Data compression of the ECG using neural network for digital holter monitor. *IEEE Eng. Med. Biol. Mag.*, 9(3), 53–57.
- [19] Al-Shrouf A., Abo-Zahhad M., and Ahmed S.M. 2003. A novel compression algorithm for electrocardiogram signals based on the linear prediction of the wavelet coefficients. *Digit. Signal Process.*, 13(4), 604–622.
- [20] Mukhopadhyay S.K., Mitra S. and Mitra M. 2011. A lossless ECG data compression technique using ASCII character encoding. *comput. Electr. Eng.*, 37(4), 486–497.
- [21] Mukhopadhyay S.K., Mitra S., and Mitra M. 2012. An ECG signal compression technique using ASCII character encoding. *Measurement*, 45(6), 1651–1660.
- [22] Mukhopadhyay S.K., Mitra S. and Mitra M. 2013. ECG signal compression using ASCII character encoding and transmission via SMS. *Biomed. Signal Process. control*, 8(4), 354–363.
- [23] Singh B., Sharma D., Singh M. and Singh D. 2014. An improved ASCII character encoding method for lossless ECG compression. *Advances in Biomedical Science and Eng.*, 1(2), 1-11.
- [24] Jalaleddine M.S., Hutchens c.G., Strattan R.D., and coberly W.A. 1990. ECG data compression techniques- A unified approach. *IEEE Trans. Biomed. Eng.*, 37(4), 329–343.
- [25] Mueller W.c. 1978. Arrhythmia detection program for an ambulatory ECG monitor. *Biomed. Sci. Instrument.*, 14, 81-85.