Electrocardiogram
Signal Compression and Decompression

Thesis submitted in partial fulfillment

for the award of the degree of

Bachelor of Technology

In

Electronics and Instrumentation Engineering

Submitted by

JALAJ CHATURVEDI [111EI0242]

Under the guidance of

Dr. Samit Ari

Assistant Professor

Department of Electronics & Communication Engineering

National Institute of Technology, Rourkela

Odisha-769008

INDIA

DECLARATION

I hereby declare that the work presented in the thesis entitled "***Electrocardiogram Signal Compression and Decompression***" is a bonafide record of the research work done by me under the supervision of Dr. Samit Ari, Department of Electronics & Communication Engineering, National Institute of Technology, Rourkela, India and that this thesis work has not been presented for the award of any other degree.

JALAJ CHATURVEDI

Dept. of Electronics & Communication Engg.

National Institute of Technology, Rourkela

Odisha-769 008

INDIA

National Institute of Technology

Rourkela, Odisha-769008, India

CERTIFICATE

Certified that this project thesis on "**ELECTROCARDIOGRAM SIGNAL COMPRESSION AND DECOMPRESSION"** is a bonafide work of "**JALAJ CHATURVEDI**" who carried out the research project under my supervision and guidance during Aug 2014-May 2015 (7th & 8th Semester). This thesis has not been submitted for any degree or academic award elsewhere.

Place:

Dr. Samit Ari

Date:

Assistant Professor

Dept. of Electronics & Communication Engg.

National Institute of Technology

Rourkela, India-769 008

ACKNOWLEDGEMENT

I would like to express my deep sense of gratitude to my supervisor **Dr. Samit Ari**, Assistant Professor, Department of Electronics & Communication Engineering, National Institute of Technology, Rourkela, for his persistent encouragement, continuous monitoring and supervision throughout the one year for this research work. I am highly indebted to him for guiding me in shaping the problem statement and providing the perception I required towards the solution. I owe to him as he endured me during his working hours and was always there for me whenever I got curious. I would also like to convey my sincere thanks to **Prof. K.K.Mohapatra**, HOD (Department of Electronics and Communication Engineering, NIT Rourkela) who found me worthy enough to undertake this project.

I would like to thank all faculty members and staffs of the Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela for their valued help throughout the project year.

I would like to comprehend by expressing my deepest thankfulness to my parents, and all near and dear. The project would not have been made possible without their blessings and ethical support.

7th May, 2015 Jalaj Chaturvedi [110EI0242]

ABSTRACT

Electrocardiogram is a method of recording the heartbeat of a patient electronically. Quicker transmission of electrocardiogram signals through channels like across wireless networks, telephone networks, hospital networks is not possible without compressing them. Here I implement an algorithm [13] for ECG (electrocardiogram) signal compression based on Delta Coding and LZW (Lempel–Ziv–Welch) encoding. To carry out compression, real time ECG data has been recorded through EKG sensors and Labview. Since the recorded signal was analog it was first converted to digital by applying Savitzky Golay Filter, uniform sampler and quantizer sequentially. Then the signal was compressed using delta coding and LZW coding. Then the compressed signal was decompressed and then further converted to analog. Finally the reconstructed signal was compared with the original signal. This was done for six different analog samples. The compression ratios and the reconstruction errors for all were calculated. The calculated compression ratios range from 1:10.51 to 1:12.56 and PRDs range from 3.33% to 4.79%. Mean compression ratio came out to be 1:11.36 and mean PRD came out to be 4.18%. The number of quantization levels was kept sufficiently large in order to reduce the quantization error zero levels. In the preprocessing stage the signal was reduced to about one-ninth of the original size while in the encoding stage it was reduced to one third. When applied sequentially, the compression ratios got multiplied.

CONTENTS

ACKNOWLEDGEMENT 4

ABSTRACT **5**

CONTENTS 6

LIST OF TABLES AND FIGURES 8

1. INTRODUCTION **9**

1.1 ECG SIGNAL 10

1.1.1 HISTORY 10

1.1.2 FUNDAMENTALS 11

1.2 NEED OF COMPRESSION 16

1.3 DATABASE 17

1.4 OBJECTIVE 17

1.5 DEFINITIONS AND FORMULAE USED 18

1.6 OVERVIEW 20

2. COMPRESSION OF ELECTROCARDIOGRAM SIGNAL 21

2.1OVERVIEW 22

2.2 NEED OF FILTERING 23

2.3 SAVITZKY – GOLAY FILTER 24

2.4 UNIFORM SAMPLING 25

2.5 QUANTIZATION 27

2.6 COMPRESSION RATIOS AND QUANTIZATION ERRORS 29

2.7 DELTA CODING 30

2.8 LZW CODING 31

2.9 RESULTS 32

2.10 CONCLUSION 32

3. DECOMPRESSION OF ELECTROCARDIOGRAM SIGNAL 33

3.1 OVERVIEW 34

3.2 LZW DECODING 35

3.3 DELTA DECODING 37

3.4 DEQUANTIZATION 38

3.5 RESULTS 39

3.6 CONCLUSION 39

4. CONCLUSION AND FUTURE WORKS 40

4.1 CONCLUSION 41

4.2 FUTURE WORKS 42

REFERENCES 43

LIST OF FIGURES

1.1	Einthoven triangle	12
2.1	Pictorial representation of the steps followed in section 2	22
2.2	Raw ECG Signal (Sample 1)	23
2.3	Savitzky-Golay Filtered ECG Signal	24
2.4	ECG Signal after Uniform Sampling	26
2.5	Quantized ECG Signal	28
2.6	Delta Coded ECG Signal	30
3.1	Pictorial representation of the steps followed in section 3	34
3.2	Signal after LZW Decoding	36
3.3	ECG Signal after Delta Decoding	37
3.4	Reconstructed ECG Signal after Dequantization	38

LIST OF TABLES

1.1	Name and position of leads in a 12 lead ECG	14
2.1	Compression Ratios and Quantization Errors after Preprocessing	29
2.2	Compilation of compression ratios after delta and LZW coding (only)	32
3.1	Final Results containing combined Compression Ratios and PRDs	39

CHAPTER 1

INTRODUCTION

Researchers have found variety of ways to find out various physiological disorders. Cardiological disorders are one of them. But despite of so many different ways and evolving technologies ECG continues to be an integral part of those techniques.

Due to this much importance of ECG , continuous researches are taking place all over the world to find out the effective methods of ECG transmission. One such method is implemented here.

1.1 ECG SIGNAL

Electrocardiography is the method of recording the electrical activity of the heart over a time period utilizing electrodes put on a patient's body. These electrodes detect the small electrical potentials generated on the skin that emerge from the heart depolarization amid every heartbeat.

1.1.1 HISTORY

The three Greek [2] words ‘electro’ , ‘kardio’ and ‘ graph’ join to form the term electrocardiograph.‘ Electro’ means electrical activity while ‘kardio’ means heart and the term graph stands for writing, and hence the term. One of the pioneers in this field is Alexander Muirhead who is believed to be the first one to attach wires to a person’s wrist to record the electrical activity of the heart.

Willem Einthoven from the Netherlands, used the string galvanometer to measure the electrical activity. It was also invented by him in 1901. The sensitivity of the galvanometer was much high than the capillary electrometer invented by Waller. The letters P, Q, R, S, and T were assigned to the various deflections. Apart from that a number of cardiovascular disorders were also described by him. He was given the Nobel Prize for Medicine in 1924 for the discovery.

1.1.2 FUNDAMENTALS

Basically there are 6 deflections [1] in a general ECG wave as stated above. Each deflection denotes a particular activity of heart. Various deflections and their corresponding activities are as under:

I. **P Wave :** It is the first deflection (positive) of ECG representing Atrial Depolarization
II. **Q Wave :** It is the second deflection (negative) of ECG representing the normal left-to-right depolarization of the interventricular septum
III. **R Wave :** It accounts for depolarization of the ventricles
IV. **S Wave :** It reflects the completion of the depolarization of the ventricles, at the base of the heart
V. **T Wave :** It represents repolarization of the ventricles
VI. **U Wave :** It's source is still unknown

There are two basic arrangements of electrodes in ECG, bipolar and unipolar leads.

Bipolar Lead: Here electrical activity of one electrode is compared to another electrode. It is by convention that a positive electrode is one in where there is a positive (upward) deflection in the ECG records when the electrical impulse flows toward it and there is a negative (downward) deflection when the impulse flows away from it.

Unipolar Lead: Here the electrical activity of one electrode is compared to a reference point that averages electrical activity and not to another electrode as is the case with the bipolar leads. This electrode is called as the *exploring electrode* and is also the positive electrode.

An imaginary formation of three limb leads in a triangular formed by the two shoulders and the pubis is known as **Einthoven's triangle.**

Given below is a figure of Einthoven triangle which briefly explains the interrelationships of the leads and the position of electrodes.

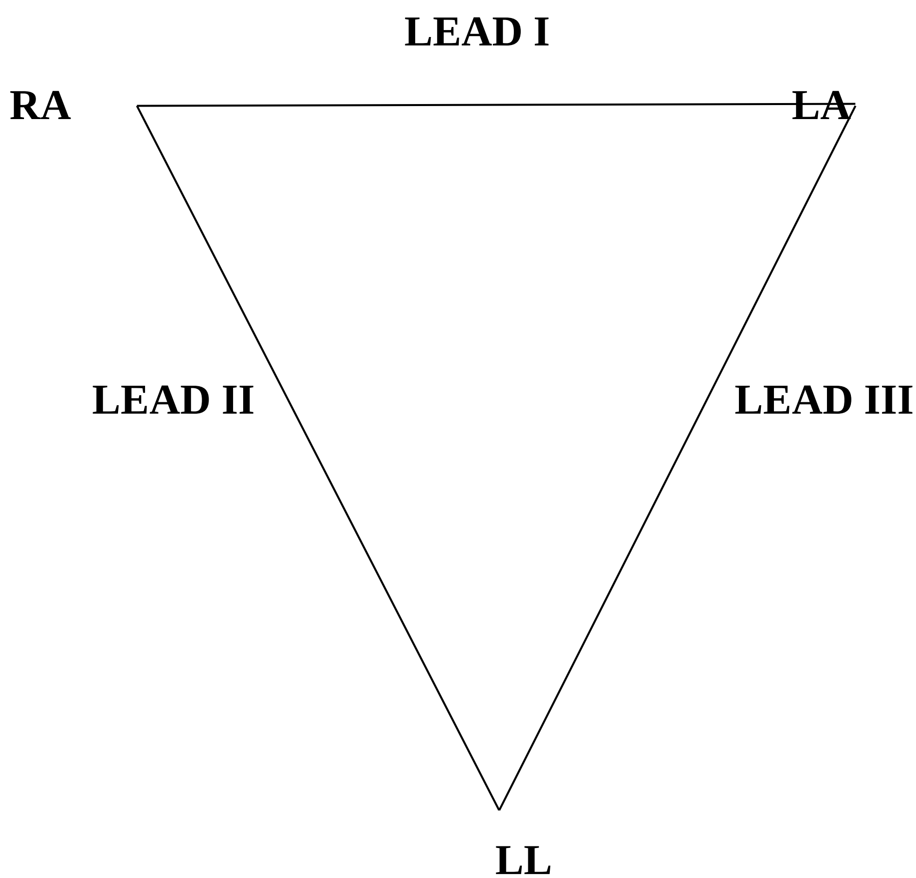

Fig. 1.1 Einthoven Triangle

Mathematically an Einthoven triangle can be expressed as :

LEAD II - LEAD I = LEAD III

An Electrocardiogram can be 1 lead, 3 lead , 6 lead or 12 lead .They are described as follows :

1) **Single lead ECG :**

Lead II is the most commonly measured vector in a single channel, one Lead ECG. As described in the Einthoven triangle, lead II is collected from the LL and RA electrodes. LL electrode site is considered positive and the RA electrode site is considered negative. Since only one lead is involved, only one amplifier module is required here.

2) **3 lead ECG**

An Einthoven triangle is just the pictorial representation of a 3 lead ECG. The electrodes have to be placed just as shown in the Einthoven triangle [5]. Since lead III is mathematically related to lead I and lead II, it can be calculated in real time. Hence there are only two independent leads, as a result of which only two amplifier modules are required.

3) **6 lead ECG**

In a 6-lead ECG, the three augmented leads (***also called Goldberger leads***) [6] are also calculated in addition to Leads I, II and III. The augmented leads are represented as: aVR, aVL, and aVF, and are also derived from Leads I, II and III. Since lead III and the augmented leads aVR, aVL and aVF [7] are mathematically related to lead I and lead II, they can be calculated in real time or after data acquisition. Hence there are only two independent leads, as a result of which only two amplifier modules are required.

4) **12 lead ECG**

In a 12 lead ECG, the number of electrodes [2] used is ten. On these electrodes is applied a thin coating of a conducting gel [11] in order to increase conductivity and decrease the physiological impedance. Here, in addition to the six limb leads and augmented leads, six precordial (chest) leads are also measured. A minimum of three amplifier modules and a Wilson Terminal are needed to measure a twelve lead ECG. The names and correct locations for each electrode are as follows:

Table 1.1 : Name and position of leads in a 12 lead ECG

ELECTRODE NAME	ELECTRODE POSITION
RA	The electrode is placed at the right arm
LA	The electrode is placed at the left arm
RL	At the right limb, lateral calf muscle
LL	On the left limb, lateral calf muscle
V_1	Between ribs 4 and 5, to the right of sterinum
V_2	Between ribs 4 and 5, to the right of sterinum
V_3	Between leads 2 and 4
V_4	Between ribs 5 and 6
V_5	Horizontally in level with V_4 in the left
V_6	Horizontally in level with V_4 in the middle

All of the leads in a 12 terminal ECG are unipolar [8] leads. Wilson's central terminal is a composite pole which is taken as the negative pole..

V_W (Wilson's central terminal) , is related to the electrodes LA, LL, and RA by the following relation:

$$\mathbf{V_W = \frac{1}{3}(V_W + V_W + V_W)}$$

V_W is the negative terminal and also the eleventh lead in a twelve lead ECG.

Lead I [9] is calculated by taking the potential difference between the LA (left arm) electrode and the RA (right arm) electrode. Mathematically, they can be represented as:

$$\mathbf{I = V_{LA} - V_{RA}}$$

Lead II [9] is calculated by taking the potential difference between the LL (left leg) electrode and the RA (right arm) electrode. Mathematically, they can be represented as:

$$\mathbf{II = V_{LL} - V_{RA}}$$

Lead III [9] is calculated by taking the potential difference between the LL (left leg) electrode and the LA (left arm) electrode .Mathematically, they can be represented as:

$$\mathbf{III = V_{LL} - V_{LA}}$$

As described above, the augmented leads are aVL, aVF, and aVR . Here negative pole is the Wilson's central terminal.

Lead aVR [10] is the potential difference between the right arm (RA) electrode and the mean of left arm (LA) and left leg (LL) electrodes. Mathematically, they can be represented as:

$$\mathbf{aVR = V_{RA} - \frac{1}{2}(V_{LA} + V_{LL}) = \frac{3}{2}(V_{RA} - V_{W})}$$

Lead aVL is the potential difference between the left arm (LA) electrode and the mean of right arm (RA) and left leg (LL) electrodes. Mathematically, they can be represented as:

$$\mathbf{aVL = V_{LA} - \frac{1}{2}(V_{RA} + V_{LL}) = \frac{3}{2}(V_{LA} - V_{W})}$$

Lead aVF is the potential difference between the left leg (LL) electrode and the mean of left arm (LA) and right arm (RA) electrodes. Mathematically, they can be represented as:

$$\mathbf{aVF = V_{LL} - \frac{1}{2}(V_{LA} + V_{RA}) = \frac{3}{2}(V_{LL} - V_{W})}$$

These days 12 lead ECG is the most commonly used ECG.

1.2 NEED OF COMPRESSION

Files those are large in size pose many problems as compared to small files for example, they fill large amounts of memory, they take more time to get transmitted to users, and they can overpower algorithms which are designed to draw results from the data which is raw. In medical applications, where high speeds of data transfer are desirable an increase in the speed of 20 times, or even 2 times is appreciable. Therefore it is clearly needed and sometimes [11] necessary to reduce the space acquired by a signal and in our case ECG, but it is equally clear that this has to be done with great care because the reduction can be lossy which in our case is not desirable. Signal compression is done because reducing the number of bits by a signal to a prescribed accuracy is the need of the day.

ECG Signal compression is a necessary tool in medical applications. Consider a case where a doctor is sitting thousands of kilometers from the patient and an immediate advice is needed.

ECG Signal transmission may take hours depending on the file size and the distance involved. In such cases signal compression can be a boon for the patient. Since the number of ECG records annually is of the order of millions and the need of sending ECG signals over telephone lines for distant analysis is increasing, an effective ECG compression technique is needed. Resources like data storage space and transmission capacity can be more efficiently used in this way.

1.3 DATABASE:

I have acquired real time analog ECG signal using the EKG sensors and recorded through labview. Six different samples were recorded and then processed and compressed. Their individual compression ratios were calculated and recorded. Then the signals were reconstructed and root mean square errors were calculated and recorded.

1.4 OBJECTIVE:

In Medical applications, signal compression is one of the most required fields of research. The main objective of my research is to implement an efficient method for ECG signal compression and decompression. In order to achieve this, the processes described below are carried out in a sequential and systematic manner. They are :

1)Preprocessing:

(i) Savitzky-Golay Filtering

(ii) Uniform Sampling at a sampling frequency of 200

(iii) 20 bit Uniform Quantization

2) Delta Coding

3) LZW Coding

4) LZW Decoding

5) Delta Decoding

6) Dequantization

7) Inverse Sampling

1.5 DEFINITIONS AND FORMULAE USED

1.5.1 The **Sampling Frequency** is defined as the reciprocal of [9] the sampling period. Mathematically the relation can be expressed as :

$$\text{Sampling Frequency} = \frac{1}{T}$$

Where,

T is the sampling period

1.5.2. **Maximum Quantization Error** is defined as the maximum round off error [12] introduced by quantization. It is given by:

$$\text{Maximum Quantization Error} = \frac{True\ Span}{2^n - 1}$$

$$Q = \frac{x_{max} - x_{min}}{2^n - 1}$$

Where,

x_{max} = Maximum value of Original Signal

x_{min} = Minimum value of Original Signal

n = bit resolution of Quantization

1.5.3. Compression Ratio may be defined as the numbers of bits required by the compressed signal to that of the original signal.. All data compression algorithms, try [13] to minimize the compression ratio by eliminating the data redundancy wherever possible. Compressed data must also represent the data with better fidelity. It is given by :

Compression Ratio $= \frac{Size\ (in\ bytes) of the\ compressed\ signal}{Size\ (in\ bytes) of the\ original\ signal}$

1.5.4 The **percentage root mean square difference (PRD) [13]** is defined as:

$$\text{PRD}\ \% = 100 \sqrt[2]{\frac{\sum_{n=1}^{N}(x(n)-\ddot{x}(n))^2}{\sum_{n=1}^{N} x^2(n)}}$$

Where,

$x(n)$ = Original Signal

$\ddot{x}(n)$ = Reconstructed Signal

1.5.5 The **normalized** version of the percentage root mean square difference [13] known as the **PRDN** is defined as :

$$\text{PRDN}\ \% = 100 \sqrt[2]{\frac{\sum_{n=1}^{N}(x(n)-\ddot{x}(n))^2}{\sum_{n=1}^{N}(x(n)-\bar{x})^2}}$$

Where,

$\bar{x}$ = arithmatic mean of x(n)

The PRD shows reconstruction reliability by comparing the data point wise. This means that it compares each and every sample to its corresponding reconstructed sample, sums the overall error and then finds the percentage error.

PRDN measures the residual root mean square error. This is yet another measure to find out the reconstruction error with the advantage that it is independent of the mean of the signal.

1.6 OVERVIEW

Chapter 1: of the thesis explains about the basics of ECG. Next a brief insight is provided about the need of signal compression. The source of and objective are presented in a lucid and simple manner. Finally the formulae and definitions used are discussed.

Chapter2: This section contains preprocessing stage of the analog ECG signal which includes filtering, sampling and quantization. In the end , compression ratios and quantization errors were calculated. Delta Coding and LZW coding are also discussed here in detail. The quantized signal has been delta coded and then LZW coded. The algorithm for LZW coding has been discussed in brief. After each step, the compression ratio was calculated and the compressed signal was displayed.

Chapter 3: Here the compressed signal has been decompressed using LZW decoding, delta decoding and Dequantization. Final results and the decompressed signal after each step were displayed.

Chapter 4 : This section contains conclusion of the overall experiment and areas of improvement are discussed.

CHAPTER 2

COMPRESSION OF ELECTROCARDIOGRAM SIGNAL

2.1 Overview:

This section includes ECG signal acquisition and preprocessing. ECG signal acquisition was done through the machine and recorded in labview. Then the signal was loaded into MATLAB and preprocessing was done using MATLAB only. Preprocessing includes filtering, sampling and quantization. Later compression ratios and Quantization Error are calculated and the results are shown.

The steps that have been followed in this section are as shown below :

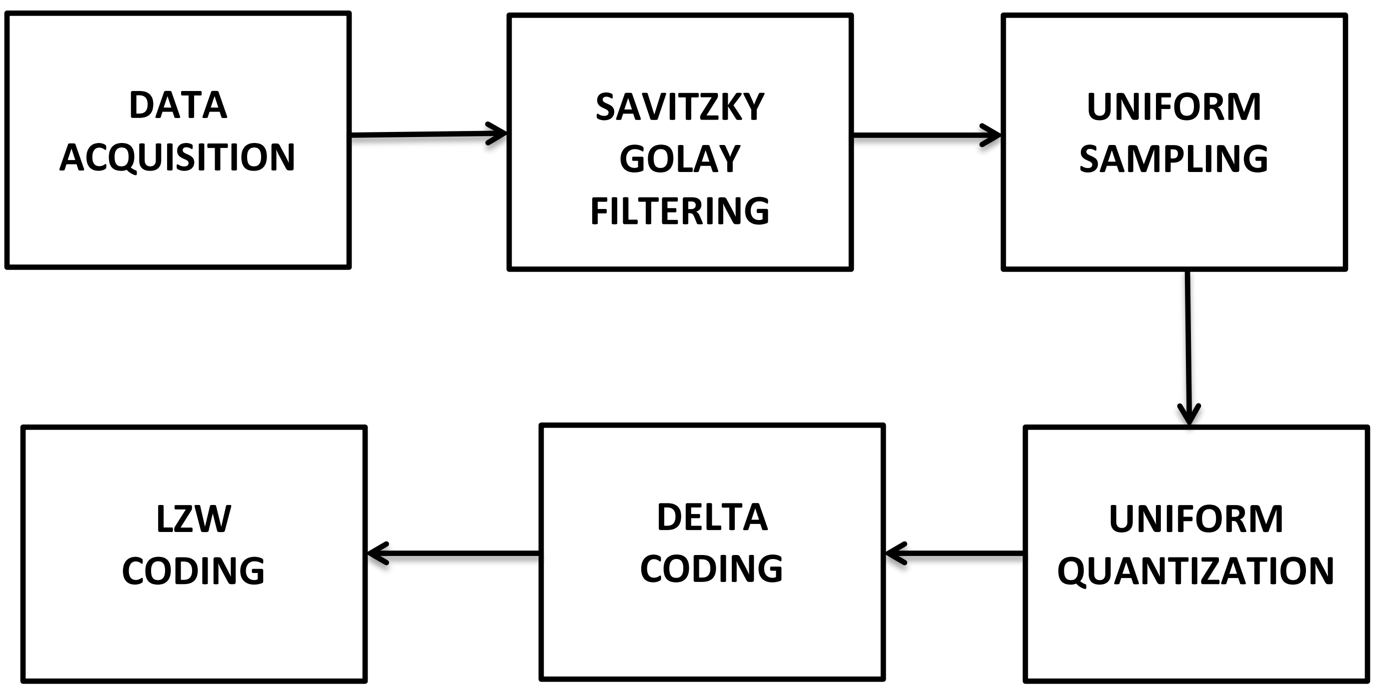

Fig. 2.1 : Pictorial representation of the steps followed in section 2

DATA ACQUISITION

EKG sensors were used to acquire analog ECG data .The Vernier EKG or ECG [14] Sensor measures cardiac electrical activity (potential waveforms produced during the contraction or expansion of the heart). With the help of Labview software the readings were ECG signal was recorded.

The following are the steps that were followed:

- The sensors were connected to the computer.
- Labview software was started on the computer.
- The program automatically identifies the sensor.
- The data was recorded.

2.2 Need of Filtering :

The raw ECG signal was applied with a filter which will not only remove unwanted noise but also decrease the size of signal thus compressing the signal.

The raw ECG signal (Sample 1) is shown below in fig 2.2

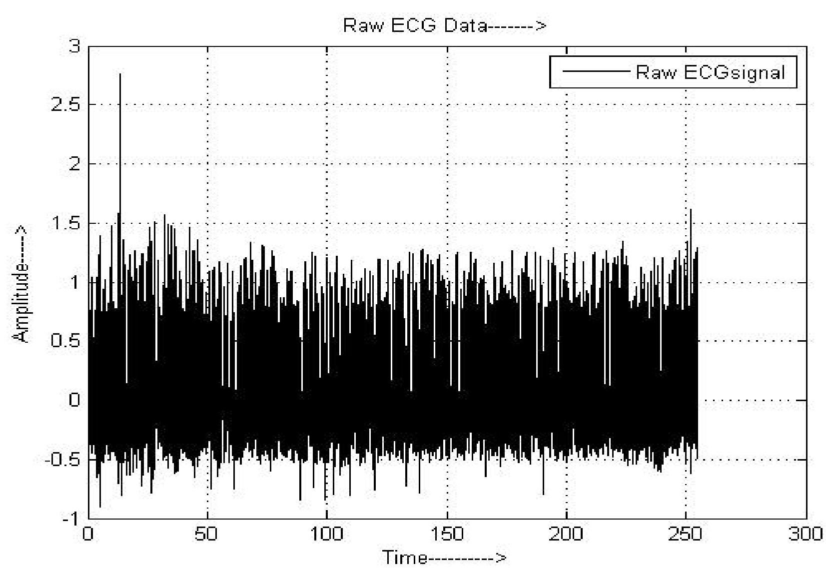

Figure 2.2: Raw ECG Signal (Sample 1)

2.3 Savitzky Golay Filtering:-

Filtering is an important pre-processing step in signal compression problems. The filtering has been done using a 6 degree Savitzky-Golay filter (SGF) and a 17-points window. The window was constant. SGFs use polynomial [18] fitting to smooth out the signal. Savitzky-Golay filters, alias. digital smoothing polynomial filters, are used to "smooth out" a signal which is noisy and whose frequency span is large. Frequency span is calculated for noiseless signals. Finite-Impulse Response (FIR) filters are good at rejecting the high frequency noise or even better than SGFs but they also reject the important parts of the signal. At this time, SGFs come into picture, whose efficiency in preserving the high frequency components of the signal cannot be questioned. After testing for various degrees of the polynomial and several dimensions of the window, the values of filter parameters were empirically decided. It has been noticed that a polynomial with smaller degrees and a window with higher dimensions lead to amplitude distortion of the R-waves while a polynomial with higher degrees and a window with smaller dimensions lead to insignificant filtering. The SG filtered ECG signal is as shown below in figure 2.3:

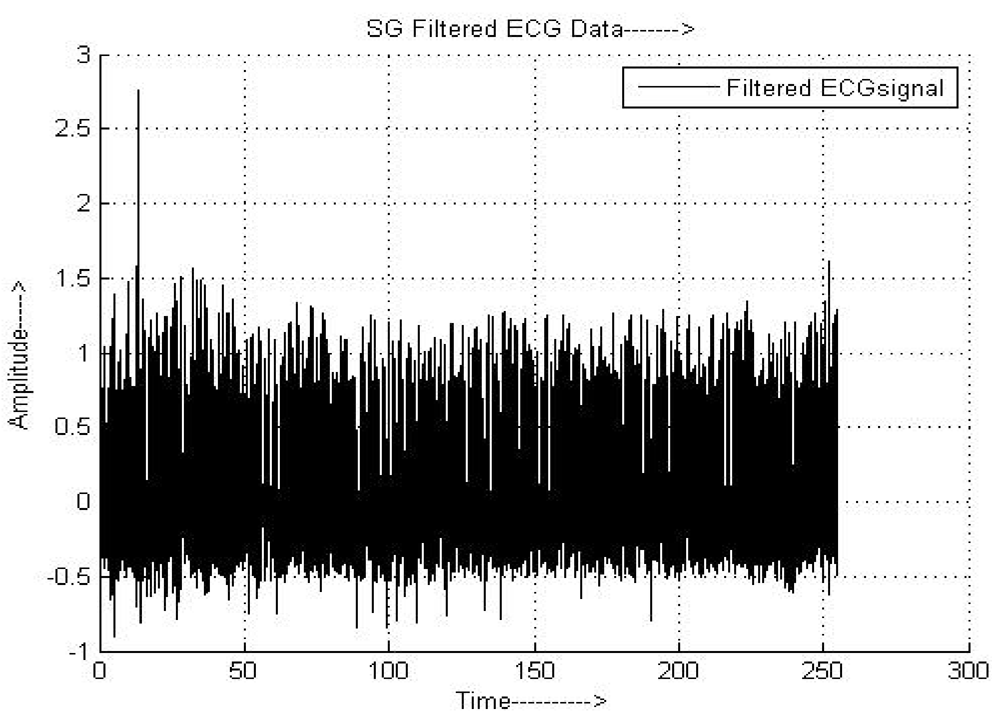

Fig. 2.3: Savitzky-Golay Filtered ECG Signal

2.4 Uniform Sampling:

Sampling is a very important step in analog to digital conversion and hence in preprocessing too. Sampling may be defined as the process of reducing a continuous signal into a discrete signal. Sampling can be done in many ways, the most common of which is uniform sampling which has been used here in this algorithm. There are other methods of sampling too like the non-uniform sampling. Non uniform sampling consists of aperiodic sampling of the analog signal. Here the number of samples is more in more important parts of the signal and less in the parts which contain noise. This is one another way of noise cancelling in signals. Though it is more efficient in noise reduction it has its own disadvantages. One of the major disadvantages is that the complexity is much more here and it becomes increasingly difficult to inverse sample the signal at the time of digital to analog conversion. Hence, uniform sampling has been used here.

Uniform sampling [21] can be mathematically represented as:

$$X(n) = X_a(nT) \qquad -\infty < n < \infty$$

Where X(n) is discrete time signal obtained after sampling of the analog signal X_a after every T seconds. The time period between successive intervals T is known as the sampling interval or the sampling period. The reciprocal of this sampling period T is defined as the sampling Frequency. Mathematically they can be rewritten as :

$$F_S = \frac{1}{T}$$

A relationship can be derived in between the continuous time variable 't' and the discrete time variable 'n' .They are related to each other according to the following relationship :

$$t = nT = \frac{n}{F_S}$$

Uniform sampling has been used here at a sampling frequency of F_s = 200 Hz. Sampling Interval is T = 0.005s.

The signal after uniform sampling is as shown below in figure 2.4

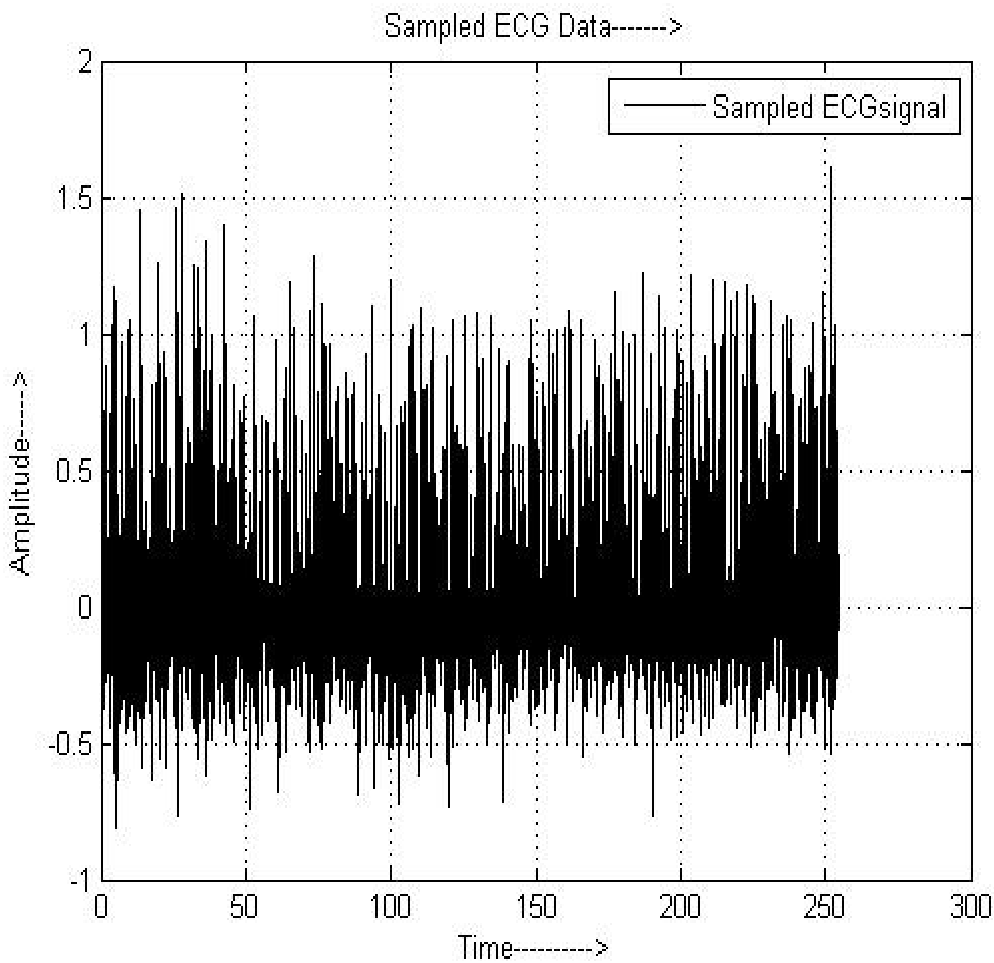

Fig. 2.4: ECG Signal after Uniform Sampling

Initially, the acquired ECG signal was analog. After sampling at a sampling frequency

F_S = 200 Hz , the number of samples obtained was 51000. The next step of analog to digital

conversion i.e. quantization has been applied in the next section.

2.5 Uniform Quantization:

Quantization is the process of converting a large set of input values to a smaller set which generally consists of rounding off values to a predefined level of accuracy. A **Quantizer** is an algorithmic function or sometimes a device that quantizes a signal. But this rounding off produces an error which is known as quantization error. [19] To be more specific the subtraction of the actual analog value and quantized digital value is known as **quantization error.** Sometimes it is also called by the name **quantization distortion**. Quantization is an integral part of analog to digital conversion or preprocessing. Quantization lies in the core of essentially all lossy compression algorithms.

Because quantization consists of rounding off the signal, it is obviously a non-linear and irreversible process (i.e., since multiple input values have the same output, generally it is impossible to recover the exact input value from the output value).

Quantization is opposite to sampling. Quantization is done on dependent variable. When we are quantizing a signal, we are actually dividing a signal into quanta (partitions). Quantization may be uniform or non-uniform i.e. if the quantization levels are uniformly spaced, then the quantization is uniform otherwise it is non-uniform.

The number of bits taken by [21] each sample of the signal is known as **bit resolution**. The number of bits also depend on the number of quantization levels used. To be more specific if the bit resolution is ‘n’ then the number of quantization levels will be 2^n .

20 bit uniform quantization was done as the final step of analog to digital conversion (the preprocessing stage).

As defined above , maximum quantization error is given by :

$$Q = \frac{x_{max} - x_{min}}{L-1}$$

Here the quantity $x_{max} - x_{min}$ can be called as the dynamic range [22].

Keeping the dynamic range fixed, if the number of quantization levels 'L' is increased, the maximum quantization error decreases. Therefore, if we keep the number of quantization levels 'L' sufficiently large, we can make the maximum quantization error negligibly less.

The signal after uniform quantization is as shown below in figure 2.5:

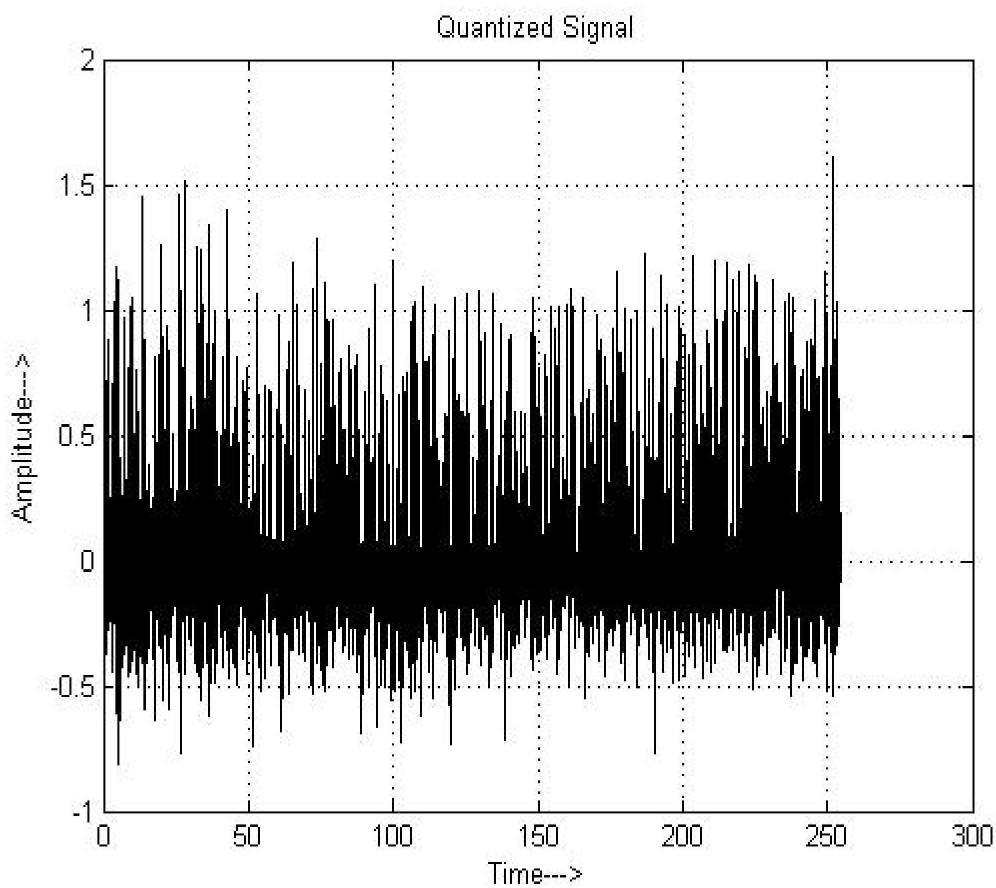

Fig. 2.5: Quantized ECG Signal

2.6 Compression Ratios and Quantization Errors:

The compression ratio obtained after the application of above steps and the quantization error are displayed below :

Table 2.1: Compression Ratios and Quantization Errors after Preprocessing

Serial Number	Sample Name	Quantization Error	Compression Ratio
1	Sample 1	2.311×10^{-6}	1 :3.28
2	Sample 2	2.562×10^{-6}	1 : 4.19
3	Sample 3	2.518×10^{-6}	1 : 3.58
4	Sample 4	2.625×10^{-6}	1 : 3.56
5	Sample 5	1.807×10^{-6}	1 : 3.92
6	Sample 6	2.843×10^{-6}	1 : 3.17

The above table shows that analog signals need much more storage space than their digital counterparts. After analog to digital conversion in the above steps, compression ratios as high as 1 : 3.61 are obtained. The table also emphasizes the fact that if the quantization levels are intelligently chosen, the quantization errors can be minimized to negligible amounts.

2.7 DELTA CODING

The Greek letter delta (Δ) is used to denote the change in a variable. The term delta encoding, refers to a technique where data is stored as the difference between successive samples instead of directly storing the samples. The first value in a delta encoded file is the same as the first value of the original data. All the following values in the encoded file are equal to the difference (delta) between the corresponding value and the previous value in the input file.

Delta encoding is typically used where the original data is smooth i.e. the differences between the adjacent values is small. The key feature of delta encoding is that it reduces the amplitudes to great extents if the original signal is smooth. In other words, it increases [24] the probability of sample's value being close to zero, and decreases the probability of it being far from zero. This uneven probability is just the thing that LZW coding needs to operate.

The signal after Delta Coding is as shown below in figure 2.6:

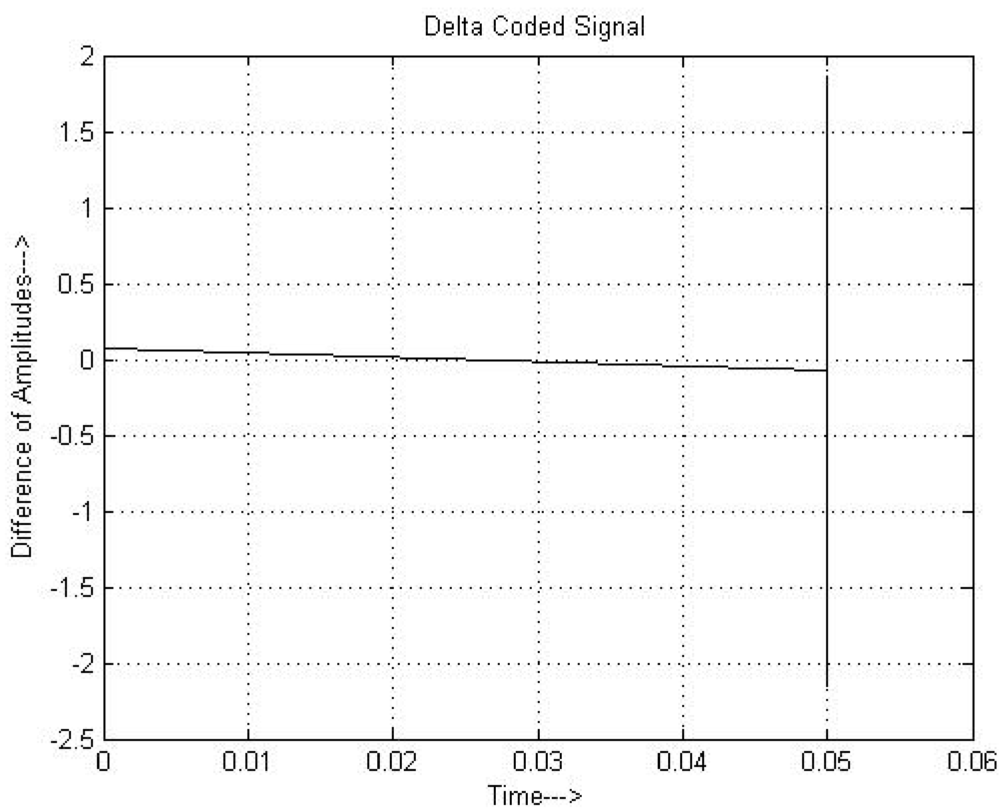

Fig. 2.6: Delta Coded ECG Signal

2.8 LZW Coding

Lempel- Ziv-Welch (LZW) is a universal lossless data compression algorithm created by Abraham Lempel, Jacob Ziv, and Terry Welch. Published by Welch in 1984, it is an improved version of the LZ78 algorithm published by Lempel and Ziv in 1978. It is a substitution or dictionary-based coding algorithm. This method reads the samples and encodes them through the creation of a dictionary of individual or sets of symbols [13]. Here the samples were first converted into strings and then numerical codes were assigned to them thus creating a dictionary of strings and numerical codes.

The algorithm involved in LZW encoding is as follows:

1. **Convert the quantized samples into strings**
2. **Initialize Dictionary with the {0,1,2,3,4,5,6,7,8,9,e,+,-} set but as character strings and their corresponding ASCII codes;**
3. **P = first input character;**
4. **Code =256**
5. **while(not end of character stream)**

```
{
   C = next input character;
   If (P * C exists in the Dictionary)   % P*C stands for P concatenate C
   P = P * C
   Else
   {
    Output: the code for P; insertInDictionary( (Code , P * C) ) ;
    Code++;
   P= C;
   }
}
```

6. **Output: the code for P;**

2.9 Results

Table 2.2: Compilation of compression ratios after delta and LZW coding (only)

Serial Number	Samples	Compression ratio
1	Sample 1	3.48
2	Sample 2	2.52
3	Sample 3	3.50
4	Sample 4	3.17
5	Sample 5	3.01
6	Sample 6	3.31

2.10 Conclusions

In this preprocessing stage, the analog ECG signal was converted to its digital counterpart. Compression Ratios as high as 1:3.61 were obtained. But due to quantization some error was also introduced in the signal which is listed above.

Six different analog ECG signals were Delta Coded and LZW Coded and their compression ratios were calculated. Mean compression ratio came out to be equal to 3.16.
This shows that data redundancy is less as compared to data loss while converting the signal from analog to digital.

CHAPTER-3

DECOMPRESSION OF ELECTROCARDIOGRAM SIGNAL

3.1 Overview

Decompression is another necessary step of the compression algorithm. Any compression algorithm is incomplete without its corresponding decompression method. Here, the decompression algorithm for the above compression method is discussed in detail. At first , the compressed signal was LZW decoded whose algorithm is given below. LZW decoding consists of finding out the decoded strings from the dictionary. Then the sequence was delta decoded. In this section, the consecutive deltas were added to each other. Finally Dequantization was done using polynomial interpolation.

The following is a pictorial representation of the steps followed in this section :

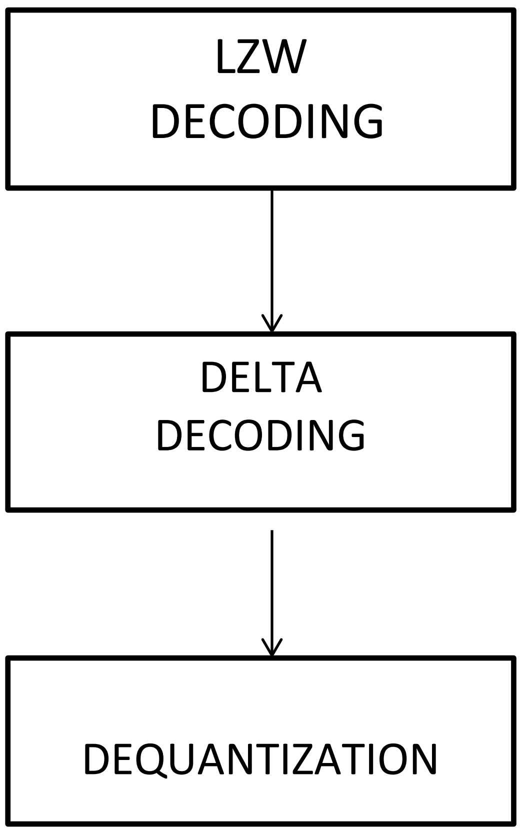

Fig. 3.1 : Pictorial representation of the steps followed in this section

3.2 LZW Decoding

LZW Decoding simply consists of taking the codes one by one from the encoded file, referring the constructed dictionary and decoding them by writing down their corresponding values. One additional step which has to be applied in this case is to reconvert the decoded character strings into numbers. Here the codes from the encoded file were taken one by one. Each code was compared with the dictionary. Then if the code lies in the dictionary, then its corresponding decoded string was stored in a separate array. The process was repeated till the end of the encoded file. Then from the array containing lengths of numbers, each string was separated and then the string was converted into numbers.

The algorithm involved during this step is as follows :

1. **Load the dictionary constructed during LZW encoding and array containing lengths of strings;**
2. **while(not end of encoded stream)**

   {

   P = input code;
3. **for i=1: length (Dictionary)**

   {

   Compare (P , Dictionary(: , i))
4. **Q= decoded value of P;**
5. **Display Q**

   }
6. **Convert the string outputs to numbers;**

The signal after LZW Decoding is as shown below in figure 3.2:

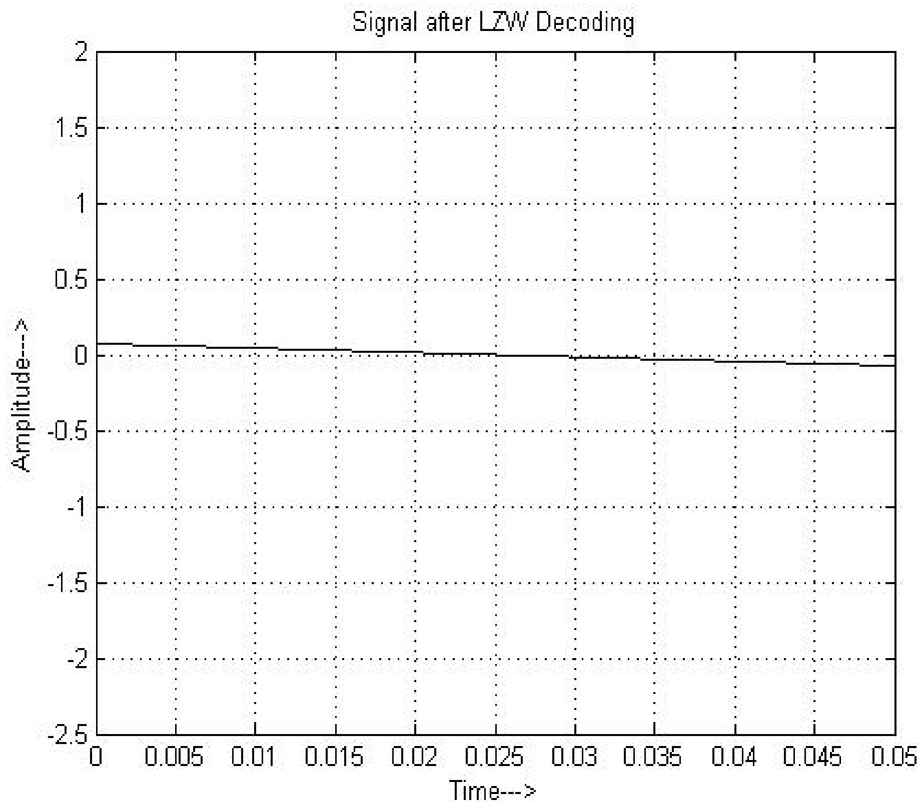

Fig. 3.2: Signal after LZW Decoding

The above figure is similar to the delta coded signal as shown above. This clearly validates the LZW coding involved. Also the curve is nearly a straight line which is because of the uniform sampling. Had the sampling been non uniform, the curve would not have been like this.

3.2 Delta Decoding

As already stated Delta Coding [26] involves storing differences of consecutive samples rather than storing the samples themselves. Hence, it becomes fairly obvious that in order to recover the original signal from the delta encoded signal consecutive deltas need to be added [15].

In this section the signal is delta decoded and the output is displayed.

The signal after Delta Decoding is as shown below in figure 3.3:

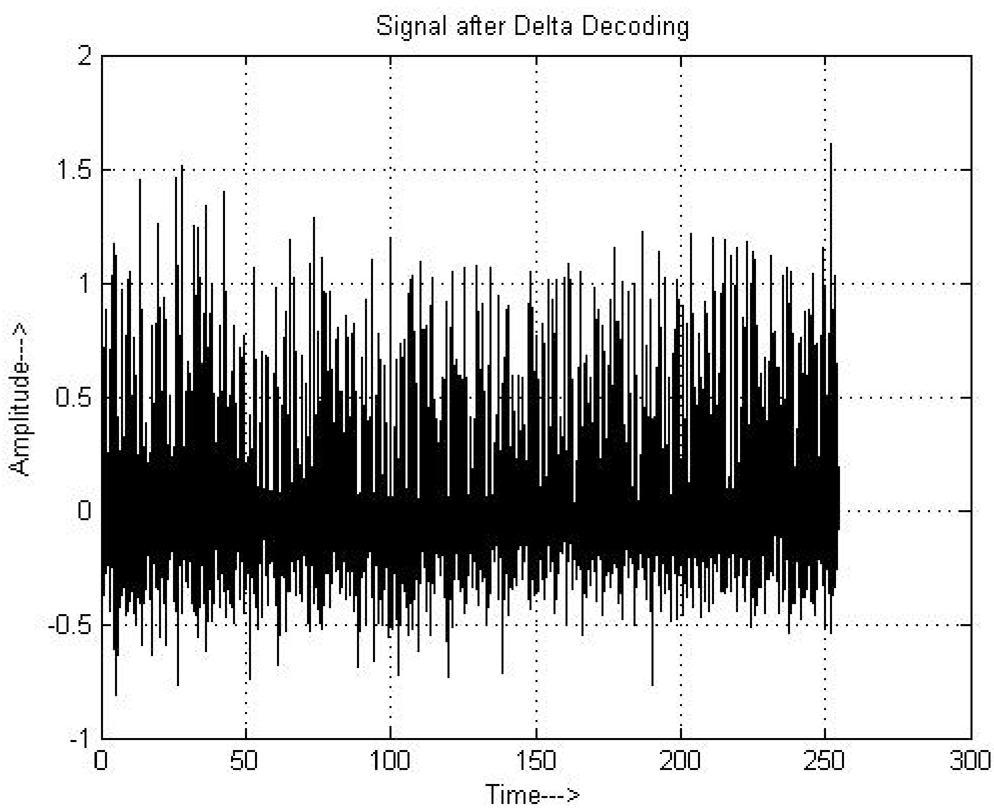

Fig. 3.3: ECG Signal after Delta Decoding

3.3 Dequantization

Dequantization [23] is an integral part of digital to analog conversion. Here it has been done using polynomial interpolation. The delta decoded values are still in digital form. In order to make them analog they need to be dequantized i.e. these discrete values need to be interconnected using interpolation [21]. Lagrange devised a method to interpolate using polynomials. In this method, if information about n points is known then a polynomial of degree n-1 is constructed. Using this polynomial, the information about the unknown points is calculated.

The signal after Dequantization is as shown below in figure 3.4:

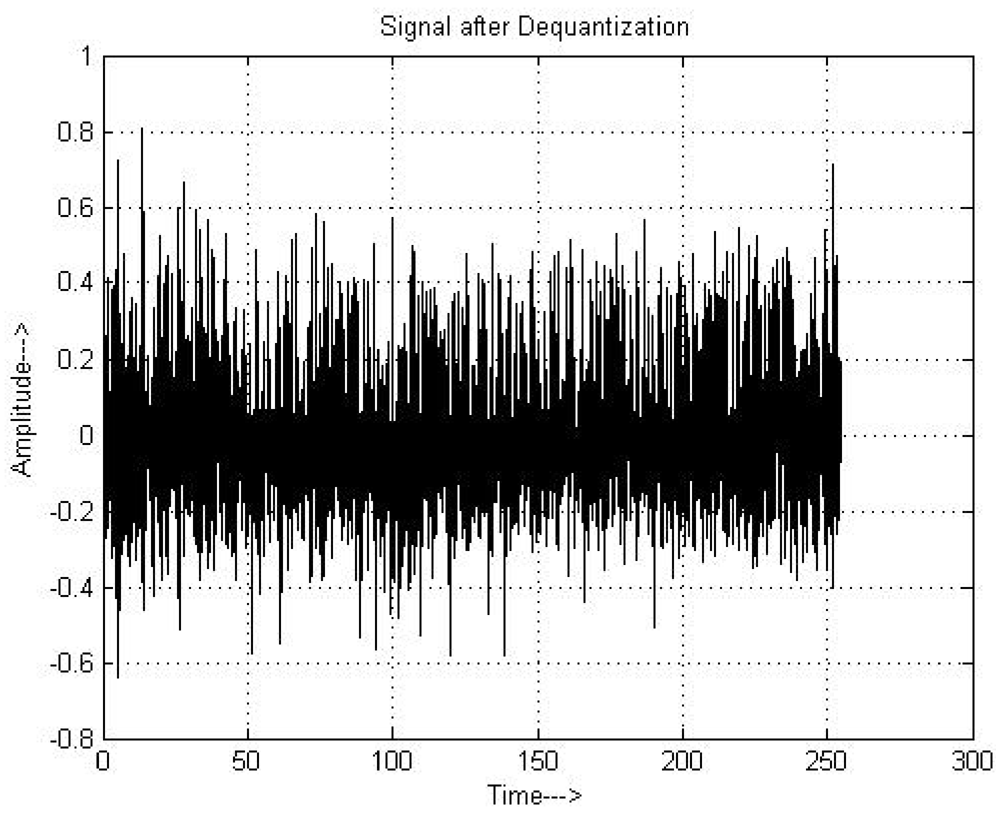

Fig. 3.4 : Reconstructed ECG Signal after Dequantization

3.4 Results

For ECG signals , PRD is a common measure to find out signal reconstruction errors. PRD and compression ratios (after sequential application of all steps) for six different analog signals was calculated and the results are displayed below:

Table 3.1 : Final Results containing combined Compression Ratios and PRDs

Serial Number	Sample Names	Compression Ratio	PRD %
1	Sample 1	1 : 11.43	4.51 %
2	Sample 2	1 : 10.567	3.33 %
3	Sample 3	1 : 12.56	4.46 %
4	Sample 4	1 : 11.29	3.58 %
5	Sample 5	1 : 11.81	4.79 %
6	Sample 6	1 : 10.51	4.41 %

3.5 CONCLUSION

A mean compression ratio of about 1 : 11.36 is obtained i.e. the efficiency of the whole compression algorithm is about one by thirty. One thing which can be concluded from the table is that on sequential application of different steps, the compression ratios of each step get multiplied to give the net compression ratio. The process of DAC conversion is lossy. Since interpolation is involved in this process, the reconstruction errors obtained are high.

CHAPTER 4

CONCLUSIONS AND FUTURE WORKS

4.1 CONCLUSION

A lot of research has been done in recent years on ECG Signal Compression. The aim of this project was to implement an efficient compression algorithm. An algorithm [13] for ECG signal compression using preprocessing, Delta Coding and LZW coding has been implemented. The algorithm [13] was tested for the compression of six different analog ECG signals recorded through EKG sensors and their compression ratios after each step was calculated. The compressed signal was reconstructed and reconstruction errors (PRD)were calculated. The mean value of the CR for the 6 records analyzed was 11.36 and the mean value of PRD was 4.18%.

The conclusions drawn after each stage are as under :

- The table of compression ratios obtained after the preprocessing stage clearly conveys the message that the analog signals need a lot more storage space than their digital counterparts. Compression ratios of about 1:3 were obtained then which shows that the digital signal uses 1/3 of the storage space compared to the analog signal in this case.

- One another thing that can be concluded from the table 2.1 is that if the bit resolution for quantization is kept high, the quantization error can be reduced near to zero levels.

- The table 2.2 gives the combined efficiency of the Delta coding and the LZW coding. The mean compression ratio obtained after delta encoding and LZW encoding is 3.16 i.e. the delta coding and LZW coding were able to compress the signal by about one third.

- Table 3.1 gives the combined compression ratios of all the steps sequentially applied. A mean compression ratio of about 1 : 11.36 was obtained i.e. the efficiency of the whole compression algorithm is about one by twelve. One thing which can be concluded from the table is that on sequential application of different steps, the compression ratios of each step get multiplied to give the net compression ratio.

- One another thing which may be concluded from Table 3.1 is that the process of DAC conversion is lossy. Since interpolation is involved in this process, the reconstruction errors obtained are so high.

4.2 FUTURE WORKS

The algorithm can be further improved in terms of compression ratios and reconstruction errors. ECG signals are of utmost importance and can be a matter of life and death for anyone, hence loss of data should be negligibly small. Hence, their lies immense possibilities in terms of PRD improvement.

Since , faster signal transmission and especially in the medical fields will always be a need, the possibilities and scopes for better compression algorithms will be also there. Therefore, better compression algorithms with smaller compression ratios are also needed.

REFERENCES

[01] S.M.S. Jalaaleddine, "ECG Data Compression Techniques-A Unified Approach", *IEEE Trans. Biomed Eng*, Vol.37, No. 5, 1991, pp. 329-342.

[02] American Heart Association Committee on Electrocardiography, "Recommendations for standardization of leads and specifications for instruments in electrocardiography and vector cardiography", *Circulation*, vol.52, no.1(Supply.), pp. 15-29, Aug. 1975.

[03] M. Crochemore and T. Lecroq, "Text data compression algorithms," in *Algorithms and Theory of Computation Handbook*, M. J. Atallah, Ed. Boca Raton, FL: CRC Press, 1998.

[04] Gerassimos D. Barlaas, Gerassimos P. Frangakis,Emmannuel S. Skordalakis, "Dictionary Based Coding for ECG Data Compression",*IEEE Trans.Info. Theory*,Vol IT-25,No.,pp 306-310

[05] M. N. Fira and L. Goras, "The H-wave detection with low computation complexity based on the Pan-Tompkins algorithm," in *Buletinul Institutului Politehnic Din Iasi, Tomul L (LIV) 2004*, vol. Fasc 3, Electrotehnica, Energetica, Electronica.

[06] R. Hermes, D. B. Geselowitz, and G. Oliver, "Development, distribution, use of the American Heart Association database for arrhythmia detector evaluation,"*Compute. Cardiol*.,Long Beach, CA, pp. 222- 243, 1980.

[07] P. De Chazaal, M. O'Dwayer, and R. B. Reilly, "Classifica- tion of heartbeats using ECG morphology and heartbeat interval fea- tures," *IEEE Trans. On Biomed. Eng.*, vol. 52, no. 8, pp. 1186–1190, Jul. 2004.

[08] The CSE European Working Party, "Establishment of a library for computer ECG Measurement programs," *Comput. Biomed. Res.*, vol. 18, pp.441-450, 1985.

[09] V. K. Murthyetal, "Analysis of power spectral densities of electrocardiograms," *Math. Biosci*. vol. 11, pp. 41-46, 1971.

[10] R. G. Gold, “Do we need a new standard for ecg ?,”*Brit. Heart J* ,vol. 52, pp. 101-110, 1985.

[11] G. K. Prasad and J. S. Sahambi, “ECG arrhythmias using multi resolution analysis and neural networks,” in *Proc. Convergent Technol. Asia-Pacific Region (Tencon)*, 2003, vol. 21, pp. 16.

[12] M.Hilton, "Wavelet and Wavelet Packet Compression of Electrocardiograms", *IEEE Trans. Biomed. Eng.*, Vol.44, pp. 393-399, 1998.

[13] C.M. Fira and L. Goras, “ An ECG Signals Compression Method and its validation using NNs”, *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 4, April 2008

[14] W. J. Tompkins and J. G. Webster, Eds., *Design of Microcomputer-Based Medical Instrumentation Englewood Cliffs*, NJ: Prentice- Hall, 1985.

[15] S. Osowski, L. T. Hoai, and T. Markiewicz, “*Support vector machine based expert system for reliable heartbeat recognition*

[16] Et.al ” Wavelet Compression of ECG Signals Using SPIHT Algorithm”*International Journal of Signal Processing 14, 2005*.

[17] Hoffa M, Ludwig C. 1850. *Einige neue versuche uber herzbeweguung*. Zeitschrift Rationelle Medizin, 9: 102-133

[18] Marey EJ. *Des variations electriques des muscles et du couer en particulier etudies au moyen de l'electrometre de M Lippman.* Compres Rendus Hebdomadaires des Seances de l'Acadamie des sciences 1876;82:975-977

[19] Márquez MF, Colín L, Guevara M, Iturralde P, and Hermosillo AG. *Common electrocardiographic artifacts mimicking arrhythmias in ambulatory monitoring.* Am Heart J 2002 Aug; 144(2) 176.

[20] Leonardo Vidal Batista Elmar Uwe Kurt Melcher a Luis Carlos Carvalho b , “ Compression of ECG signals by optimized quantization of discrete cosine transform coefficients”,*Medical Engineering and Physics*,2002,pp.170-189.

[21] M.L Hilton, “Wavelet and wavelet packet compression of electrocardiograms”, *IEEE Trans. Biomed. Eng.*, vol.44, pp.328-413, 1997.

[22] J.G. Proakis, D.G. Manolakis, “*Digital Signal Processing Principles, Algorithms and Applications”*, Pearson Education, pp. 19-37.

[23] M. Unser, A. Aldroubi, A review of wavelets in biomedical applications, *Proc. IEEE 84 (4) (1996)*, pp. 619–630.

[24] Shang-Gang Miaou, Heng-Lin Yen, Chih-Lung Lin, “Wavelet Based ECG Compression using Dynamic Vector Quantization with Tree Code vectors in Single Codebook”, *IEEE trans. Biomedical Eng 2002*, 49(7),pp.680-690.

[25] I.Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inform. Theory 36 (5) (1990)* , pp.970–989.