Application of LZW Technique for ECG Data Compression

Mridul Kumar Mathur [1], Dr. Akhil Ranjan Garg [2], Prof. Mukesh Upadhayay[3]

[1] DEPARTMENT OF COMPUTER SCIENCE, LACHOO MEMORIAL COLLEGE OF SCIENCE AND TECHNOLOGY, JODHPUR, RAJASTHAN, INDIA

[2]DEPARTMENT OF ELECTRICAL ENGINEERING, M.B.M ENGINEERING COLLEGE, JAI NARAIN VYAS UNIVERSITY, JODHPUR, RAJASTHAN, INDIA

[3] DEPARTMENT OF COMPUTER SCIENCE, LACHOO MEMORIAL COLLEGE OF SCIENCE AND TECHNOLOGY, JODHPUR, RAJASTHAN, INDIA

mathur_mridul@yahoo.com[1], garg_akhil@yahoo.com[2], mubit@yahoo.com[3]

ABSTRACT

The ECG is an electronic method of recording the heart beat of the patient. The characteristic of heart beat is determined by PQRST curve. By attaching additional leads while recording ECG signals they require higher sampling rates and finer amplitude resolution, for noncardiac signals such as blood pressure and respiration it requires huge amount of storage space. For the fast transmission of ECG signals across wireless network, from ambulance to hospital, telephone networks, hospital networks is not possible without compressing them.
The sampling rate of collecting ECG data is at 130 Hz. The heart beat for the sample is determined by the time intervals b/w two RR. RR interval is the time measured between two successive R peaks. The occurrence of number of intervals in a minute determines the heart rate (e.g. beats/min.)

In this paper, we have applied Lempel-Ziv-Welch (LZW) method of compression over the ECG data. To carry out the compression of ECG data we have used MIT BIH patient record number 101, 107, 112 and 117. The compression is applied over ECG data converted to binary normalized form, such that the data values after normalization ranges in ASCII set. The ratio of compression obtained for the given samples ranges from 74% to 87%. Again, it has been observed that, the reconstructed signal after decompression is similar to the original normalized signal (Lossless compression).

Index Terms— Compression ratio, rhythms, Electrocardiogram, Sampling, Frequency, Threshold, Redundancy.

1 INTRODUCTION

The cost of storage devices has reduced drastically in recent times, but to store long ECG signals like 24-Hour heart rate monitoring, a very large storage capacity is required. Again, to carry out transmission of ECG data over networks i.e., for telemedicine services the time required to transmit large files is very high. Thus, in both the cases, i.e., for the purpose of storage and for the purpose of transmission compression of ECG data is much in need. Various data compression techniques are available providing lossless and lossy compressions. The requirement of ECG signals is to apply with such compression techniques, such that, the signal should be compressed but after decompression (reconstruction), the signal must contain all the required information for the purpose of diagnosis or in other words it should be in clinically acceptable form.
Lossy compressors sometimes may hide such details, whereas lossless compressor generates the exactly same information as the original signal. Data compression techniques are evaluated not only in terms of compression ratio and compression and decompression bandwidths achieved but also based on their performance when the compressed data must be sent over any of the available wireless networks.
Conceptually, data compression for ECG signals is the process of detecting and eliminating redundancies and filtering the noncardiac signals such as blood pressure and respiration and so on from a given data set. Data compression algorithms are needed because of following reasons:

1. To increase storage capacity of ECG's as databases for future comparison or assessment.

2. Implementation of economical rapid transmission of data over wireless or wired network.

3. It is very useful when transmission is required from home or ambulance to expert doctor for suggestion for a serious patient.

The key requirement of any data compression algorithm is to preserve the significant clinical information content upon reconstruction or decompression to obtain good diagnostic quality.

2 Data Compression

Data compression methods can be classified into following methods.

(a) Lossless compression Method: in this method there is an exact reconstruction of the original signal but the low data rates are not achieved.

(b) Lossy Compression method: under this method higher compression ratios can be achieved with a certain amount of distortion from the original signal. The distortion must remain so small that it would not affect the diagnostic content of the ECG.

It is necessary to implement noise filter on the ECG signal before compression under both these methods.

LZW Method

Lempel- Ziv-Welch (LZW) is a universal lossless data compression algorithm created by Abraham Lempel, Jacob Ziv, and Terry Welch. It was published by Welch in 1984 as an improved implementation of the LZ78 algorithm published by Lempel and Ziv in 1978. The two different types of compression lossy & lossless differ in one respect: Lossy compression accepts a slight compression of data to achieve compression. Lossy compression is done on analog data stored digitally, with primary applications being graphics and sound files.

Lempel-Ziv is substitution or dictionary-based coding algorithm. This method reads strings of symbols and encodes them through the creation of a dictionary of individual or sets of symbols.

In text compression, the LZW algorithm starts with a string of characters containing binary data in the form of characters ranging from (0-255) ,then a dictionary is made for each repetitive pattern. Dictionary starts from 256 -4096 code. Then every new pattern is stored in dictionary and encoded accordingly. The compression algorithm is illustrated in Figure 1.

Decoding algorithm uses compressed file and stores the new pattern in another dictionary and decode the code matches it in the dictionary and the same code pattern is generated again and file is decompressed. Figure 2 illustrates the decoding algorithm.

```
Initialize Dictionary with 256 single character
strings and their corresponding ASCII codes;
P  := first input character;
Code ← 256
while(not end of character stream)
 {
    C := next input character;
    if(P + C exists in the Dictionary)
      P := P + C
   Else
    {
     Output: the code for P;
     insertInDictionary( (Code , P + C) ) ;
     Code++;
     P:= C;
    }
 }
  Output: the code for P;
```

Figure 1: Compression Algorithm

```
Initialize Dictionary with 256 ASCII codes
and corresponding single character strings
as their translations;
P :=first input code;
Output: string(PreviousCodeWord) ;
C := character(first input code);
Code := 256;
while(not end of code stream){
     C := next input code ;
      if(C exists in the Dictionary)
         String := string(C) ;
   else
        String :=string(P) + C ;
        Output: String;
        Char :=first character of String ;
        insertInDictionary( (Code , string(P) + C )   );
        P ← C ;
  Code := Code + 1  ;
```

Figure 2: Decoding Algorithm

3 ECG Data compression

In our experiments we have taken into consideration of records of MIT-BIH Patient number 101,107,112 and 117. The data has been drawn from reliable website www.physionet.org.

For the purpose of compression and decoding we have developed programs in Delphi language. The steps followed for the purpose are stated as:

1. First, the patient data is converted to binary format.
2. The binary format of data is followed by LPF (Low Pass Filter), where by a threshold value is defined, such that, any signal value less than the threshold value is shifted to this minimum threshold, resulting the elimination of noise and without disturbing the relevant signal components or the pattern of plot.
3. Since, from observations of data it has been found that the numeric ranges of patient data from R1 to R2. We have combined the amplitude quantification by a factor of k a small value ranging from 2 to 5. Thus, the data is normalized by a divisor factor k, such that, the data values fall in a range of 0 to 255; the ASCII character set, and subsequently each data value is available as ASCII char.
4. This data format of ECG signal is now applied with Compression Algorithm as stated in Figure 1.
5. For the purpose of decoding (decompression) the compressed data is applied with algorithm as stated in Figure 2.

4 Experimental Result

For the purpose of illustration, we here present the results of data of patient record number 101. First 80 data values are shown in Table 1. (Row wise)

Table 1: Patient Record Number 101 Raw data

976	979	967	952	949	952	959	954
975	981	964	951	949	948	955	955
973	978	962	949	950	950	952	958
973	980	962	948	952	953	951	956
977	976	960	951	948	956	952	955
980	972	961	954	949	952	955	955
981	973	958	954	951	950	958	955
980	974	954	951	950	950	957	959
981	974	954	950	952	952	957	958
978	972	954	946	952	955	953	958

The normalized data of Table-1 is shown in Table-2.

Table 2: Normalized data of Table 1

58	59	55	50	49	50	53	51
58	60	54	50	49	49	51	51
57	59	54	49	50	50	50	52
57	60	54	49	50	51	50	52
59	58	53	50	49	52	50	51
60	57	53	51	49	50	51	51
60	57	52	51	50	50	52	51
60	58	51	50	50	50	52	53
60	58	51	50	50	50	52	52
59	57	51	48	50	51	51	52

The wave plot (graphical plot) of data of Table-2 for first 1500 data values is shown in Figure 3.

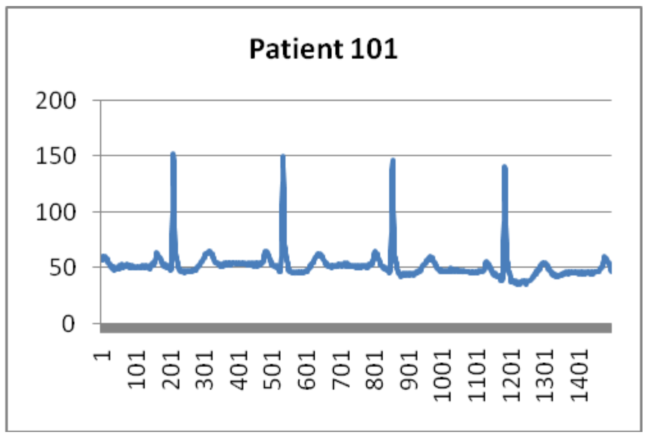

Figure 3: Wave plot of Normalized data

The data obtained after normalization is applied with LZW compression and the result of compression is stated as

1. Original File Size = 1842 K Bytes
2. Compressed File Size = 19 K Bytes
3. Code File Size = 108 K Bytes

Here it is to note that since the data is in normalized form, thus the code file is needed to be appended with the compressed file. Hence the compression ratio obtained for the selected sample is 85.49%

The compressed file is then again applied with decompression algorithm and the result of first 1500 data of decompressed output file is shown in Figure 4.

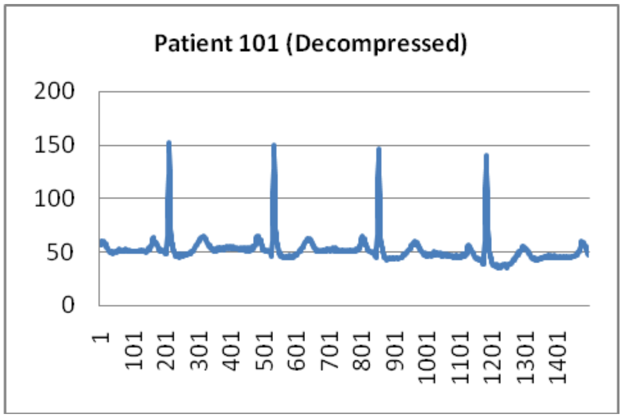

Figure 4: Decompressed data output plot

Thus, it can be concluded that the LZW compression works very well with ECG data, without any significant loss in the plot of the signal for the purpose of analysis of patient health record.

It is also here to mention that that RR interval is also same both the graphical plots.

Bibliography

[01] S.M.S. Jalaleddine, "ECG Data Compression Techniques-A Unified Approach", IEEE Trans. On Biomed Eng, Vol.37, No. 4, 1990, pp. 329-343.

[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. 2 (Supply.), pp. 11-31, Aug. 1975.

[03] Monica Negoita, Liviu Goras, on "COMRESSION ALGORITHM FOR ECG SIGNAL", Digital Signal Processing, vol. 13, 2003, pp. 604-622.

[04] Gerassimos D. Barlas,Gerassimos P. Frangakis,Emmanuel S. Skordalakis, "Dictionary Based Coding for ECG Data Compression",IEEE Trans. on Info. Theory,Vol IT-24,No.,pp 300-312

[05]American Heart Association Committee on "Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiolppogy", "Recommendations for standards of instrumentation and practice in the use of ambulatory electrocardiography," Circulation, vol. 71, 626A-636A. Mar. 1985.

[06]. R. Hermes, D. B. Geselowitz, and G. Oliver, "Development distribution, and use of the American Heart Association database for ventricular arrhythmia detector evaluation," Compute. Cardiol.,Long Beach, CA, pp. 263-266, 1980.

[07] R. G. Mark, P. S. Schluter, G. B. Moody, P. H. Devlin, and D. Chenroff, "An annotated ECG database for evaluating arrhythmia detectors," in Proc. IEEE Frontiers Eng. Health Care, 1982, pp.

[08] The CSE European Working Party, "Establishment of a reference library for evaluating computer ECG measurement programs," Comput. Biomed. Res., vol. 18, pp. 439-457, 1985.

[09] V. K. Murthyetal, "Analysis of power spectral densities of electrocardiograms," Math. Biosci. vol. 12, pp. 41-51, 1971.

[10] R. G. Gold, "Do we need a new standard for electrocardiographs?,"Brit. Heart J ,vol. 54, pp. 119-120, 1985.

[11] D. Taylor and R. Vincent, "Signal distortion in the electrocardiogram due to inadequate phase response," IEEE Trans. Biomed. Eng., vol. BME-30, pp. 352-356, June 1983.

[12] M.Hilton , "Wavelet and Wavelet Packet Compression of Electrocardiograms", IEEE Trans. Biomed. Eng., Vol.44, pp. 394-400, 1997.

[13] W. J. Tompkins and J. G. Webster, Eds., Design of Microcomputer-Based Medical Instrumentation Englewood Cliffs, NJ: Prentice-Hall, 1981.

[14] Vibha Aggarwal and Manjeet Singh "ECG compression using Wavelet Packet, Cosine Packet and Wave Atom Transforms " IEEE Transactions on Biomedical Engineering, Vol 1 , Oct 2002

[15] Mohammad Pooyan, Ali Taheri, Morteza Moazami-Goudarzi, Iman Saboori " Wavelet Compression of ECG Signals Using SPIHT Algorithm" International Journal of Signal Processing 14, 2005.

[16] D. Taylor, P. Finch, and R. Vincent, "Linear phase filtering-A new approach to distortion-free electrocardiograms," Comput. Cardiol., Linkoping, Sweden, pp. 283-286, Sept. 1985.

[17] R. L. Longini, J. P. Giolma, C. Wall, 111, and R. F. Quick, "Filtering without phase shift," IEEE Trans. Biomed. Eng., vol. BME-22, pp. 432.-433, Sept. 1975.

[18] M.Brito, J. Henriques, P. Gill and M. Antues. "A Predictive Adaptive Approach to Generic ECG Data Compression", 2005, PP.32-37.