IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 4, No 1, July 2013
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
www.IJCSI.org

New Efficient Technique for Compression of ECG Signal

Nidhal K. El Abbadi[1] Abbas M. Al-Bakry[2]
1 University of kufa
Najaf, Iraq

2 University of Babylon
Babylon, Iraq

Abstract
Data compression is a common requirement for most of the computerized applications. There are number of data compression algorithms, which are dedicated to compress different data formats. This paper examines lossless data compression algorithm for ECG data by using new method to process the ECG image strip, and compares their performance.
We confirming that the proposed strategy exhibits competitive performances compared with the most popular compressors used for ECG compression.

Key words: Data compression, ECG, compression ratio, image compression, lossy compression , lossless compression.

1. Introduction

An ECG is simply a representation of the electrical activity of the heart muscle as it changes with time, usually printed on paper for easier analysis. Like other muscles, cardiac muscle contracts in response to electrical *depolarization* of the muscle cells. It is the sum of this electrical activity, when amplified and recorded for just a few seconds that we know as an ECG.

The amplitude, or voltage of the recorded electrical signal is expressed on an ECG in the vertical dimension and is measured in millivolts (mV). On standard ECG paper 1mV is represented by a deflection of 10 mm. An increase in the amount of muscle mass, such as with left ventricular hypertrophy (LVH), usually results in a larger electrical depolarization signal, and so a larger amplitude of vertical deflection on the ECG.
An essential feature of the ECG is that the electrical activity of the heart is shown as it varies with time. In other words we can think of the ECG as a graph, plotting electrical activity on the vertical axis against time on the horizontal axis. Standard ECG paper moves at 25 mm per second during real-time recording. This means that

when looking at the printed ECG a distance of 25 mm along the horizontal axis represents 1 second in time.

ECG paper is marked with a grid of small and large squares. Each small square represents 40 milliseconds (ms) in time along the horizontal axis and each larger square contains 5 small squares, thus representing 200 ms. Standard paper speeds and square markings allow easy measurement of cardiac timing intervals. This enables calculation of heart rates and identification of abnormal electrical conduction within the heart (Figure 1).

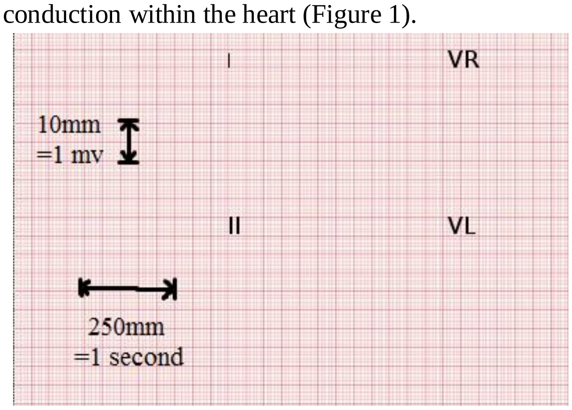


Fig 1: sample of ECG strip

Electrocardiogram (ECG) compression has been the object of numerous research works. Their main objective is to reduce the amount of digitized ECG data as much as possible with a reasonable implementation complexity while maintaining a clinically acceptable signal. Consequently, reduction of digitized ECG data allows improvement of storage capacity in the memory and/or reduces the cost of transmission.

IJCSI
www.IJCSI.org

The central goal of electrocardiogram (ECG) data compression techniques is to preserve the most useful diagnostic information while compressing a signal to an acceptable size (Al-Shrouf et al, 2003). Lossless compression is the best choice as long as the compression ratio is acceptable, but it cannot usually offer a satisfactory compression ratio (CR). To obtain significant signal compression, lossy compression is preferable to a lossless compression (AHMED et al, 2007). In this case, compression is accomplished by applying an invertible orthogonal transform to the signal, and one tries to reduce the redundancy present in the new representation. Due to its decorrelation and energy compaction properties and to the existence of efficient algorithms to compute it, discrete cosine transforms and modified discrete cosine transform have been widely investigated for ECG signal compression. Over the years, a variety of other linear transforms have been developed which include discrete Fourier transform (DFT), discrete wavelet transform (DWT) and many more, each with its own advantages and disadvantages (Daubechies, 1998).

One of the most difficult problems in ECG compression and reconstruction is defining the error criterion that measures the ability of the reconstructed signal to preserve the relevant information. As yet, there is no mathematical structure to this criterion, and all accepted error measures are still variations of the mean square error or absolute error, which are easy to compute mathematically, but are not always diagnostically relevant.

ECG signals contain a large amount of information that requires large storage space, large transmission bandwidth, and long transmission time. Therefore, it is advantageous to compress the signal by storing only the essential information needed to reconstruct the signal as in fig 2.

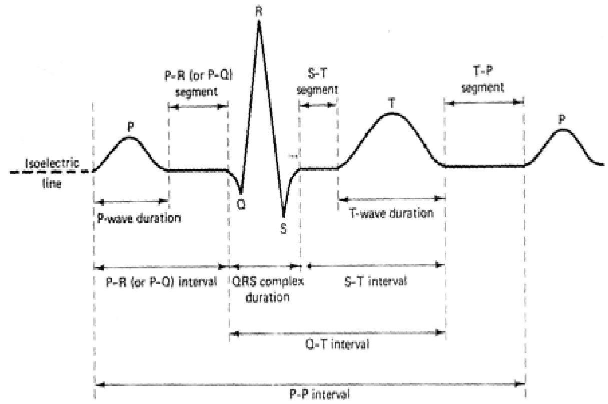


Fig 2: the essential information in ECG strips

Thus, in ECG signal compression, the objective is to represent the signal using fewer bits per sample, without losing the ability to reconstruct the signal. ECG data compression techniques are typically classified into three classes (Cardenas and Lorenzo, 1999).These classes are: direct compression, transform coding, and parameter extraction methods. In the direct compression techniques, redundancy in a data sequence is reduced by examining a successive number of neighboring samples. An example of this approach is the coordinate reduction time encoding system (CORTES). In the transform coding techniques, redundancy is reduced by applying linear transformation to the signal and then compression is applied in the transform domain rather than in the time domain. Examples of this type are Fourier transforms and wavelet transforms. In the parameter extraction techniques, the signal can be reconstructed by extracting a set of parameters from the original signal, which are used in the reconstruction process (Nave and Cohen, 1993).

This paper is organized as follow. Section 2 shows the related work. Section 3 presents an idea about the compression measures. Section 4 displays the research methodology. Finally, the paper is concluded in section 5.

2. ECG Compression Algorithms

Many existing compression algorithms have shown some success in electrocardiogram compression; however, algorithms that produce better compression ratios and less loss of data in the reconstructed data are needed. This project will provide an overview of several compression techniques and will formulate new emerging algorithms that should improve compression ratios and lessen error in the reconstructed data. Following some of these algorithms:

(Ahmed et al, 2007), present compression technique for ECG signals using the singular value decomposition (SVD) combined with discrete wavelet transform (DWT). The central idea is to transform the ECG signal to a rectangular matrix, compute the SVD, and then discard small singular values of the matrix. The resulting compressed matrix is wavelet transformed, threshold and coded to increase the compression ratio. The results showed that data reduction with high signal fidelity can thus be achieved with average data compression ratio of 25.2:1.

(Chawla, 2009), in this paper Principal Component Analysis (PCA) is used for ECG data compression, denoising and decorrelation of noisy and useful ECG components or signals signal-to-noise ratio is improved

(ALSHAMALI, 2010), this paper proposes a new wavelet-based ECG compression technique. It is based on optimized thresholds to determine significant wavelet coefficients and an efficient coding for their positions. Huffman encoding is used to enhance the compression ratio.

(Bendifallah et al, 2011), An improvement of a discrete cosine transform (DCT)-based method for electrocardiogram (ECG) compression is presented. The appropriate use of a block based DCT associated to a uniform scalar dead zone quantizes and arithmetic coding show very good results.

(Anubhuti et al, 2011), A wide range of compression techniques based on different transformation techniques like DCT, FFT; DST & DCT2 were evaluated to find an optimal compression strategy for ECG data compression. Wavelet compression techniques were found to be optimal in terms of compression.

(ALSHAMALI, 2011), adaptive threshold mechanism to determine the significant wavelet coefficients of an electrocardiogram (ECG) signal is proposed. It is based on estimating thresholds for different sub-bands using the concept of energy packing efficiency (EPE). Then thresholds are optimized using the particle swarm optimization (PSO) algorithm to achieve a target compression ratio with minimum distortion.

3. Compression measures

The size of compression is often measured by CR, which is defined as the ratio between the bit rate of the original signal (boriginal) and the bit rate of the compressed one (bcompressed) (Jalaleddine et al, 1990).

$$\text{Compression Ratio} = \frac{\text{Size after compression}}{\text{Size before compression}}$$

$$\text{Compression Factor} = \frac{\text{size before compression}}{\text{size after compression}}$$

$$\text{saving percentage} = \frac{\text{size before compression - size after comression}}{\text{size before compression}}$$

The problem of using the above definition of CR is that every algorithm is fed with an ECG signal that has a different sampling frequency and a different number of quantization levels; thus, the bit rate of the original signal is not standard. Some attempts were made in the past to define standards for sampling frequency and quantization, but these standards were not implemented, and developers of the algorithms still use rates and quantizes that are convenient to them. The number of bits transmitted per sample of the compressed signal has been used as a measure of information rate. This measure removes the dependency on the quantize resolution, but the dependence on the sampling frequency remains. Another way is to use the number of bits transmitted per second as a compression measure. This measure removes the dependence on the quantizes resolution as well as the dependence on the sampling frequency.

4. Proposed Algorithm

There are many different devices used for ECG, all shared to provide data to physicians to help them analyze ECG data to detect abnormalities, all of these devices draw ECG waveform on the specific paper, the physician can read the ECG strip to decide whether the heart normal or not, and determined if there is an imbalance or diseases facing the patent heart.

In the proposed method, the ECG data used in the image form, and treated as image data, then compressed by three stages. It is very easy to get digital image from ECG devices.

In general, the ECG image (strip) consists from the image background (baselines) and the ECG waveform draws on this strip according to heart activity (essential information).

The baselines are standard, and the distance between the lines are fixed as explained in previous section.

This work aims to isolate the ECG waveform data from background (baselines).

It is clear that the ECG waveform represent the useful data needed for physicians, as opposite of baselines which represent an assistance shape help to interpret the ECG waveform which change according to patent status. Usually the ECG waveform generally draws on baselines with dark color.

The research focused on possibility to isolate ECG waveform data from baselines, and retrieve it's later without loss data (lossless or almost lossless compression).

Isolation of ECG waveform data from baselines data not easy work, due to interference between ECG waveform data and baselines data, this causes either to lost some of ECG waveform data or save some of baselines data with ECG waveform data and make it noisy data, both cases confuse the physician and not help him in diagnosis.

ECG image data (strip) represent as a matrix of pixels, each pixel consist of three bytes (one for red color, other for green color and the last one for blue color), it is possible to imagine these matrix as three channels (channel for each base color).

The experiment proves that the isolation of ECG waveform from the origin image not useful due to lose a lot of data in addition to noise, same thing happened when the origin image converted to gray scale image.

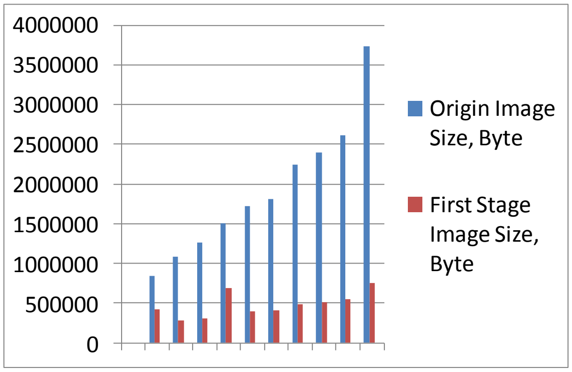


Fig 3: comparing compression ratio in first and second stage

The best way to isolate the ECG waveform data done by divide the origin image data to three sub-images data, each for one color channel, then process the red sub-image to isolate the ECG waveform data, and neglect the other two sub-images, this step reduce size to one third of origin size.

The red sub-image processed to isolate ECG waveform data by filtering followed by applying Sobel edge detection algorithm. This step will reduce the image size to more than 80% of origin size. The Sobel edge detection algorithm is the best algorithm among the other edge detection algorithms for this work to process red channel image.

The result of this stage is binary image, this produced by converting the back ground color (baselines to black color) and the ECG waveform color converted to white color (black and white image need only one bit to represent it's).

The last step is to compress the binary image by using DEFLATE algorithm which is lossless compression algorithm.

The result of average compression percent from applying the proposed algorithm on 9 different ECG images was 99.125% as shown in fig 4.

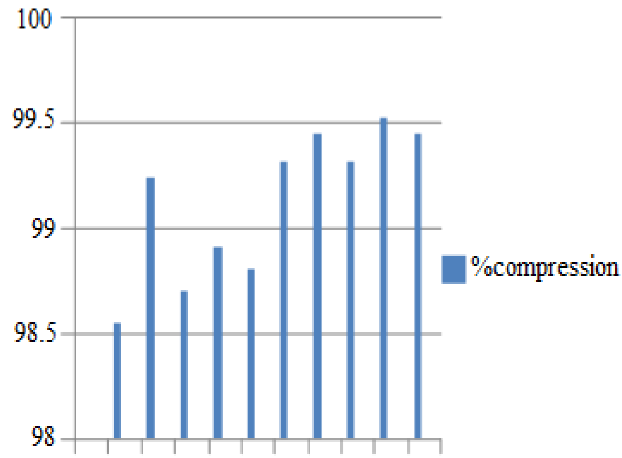


Fig 4 compression rate for 9 different ECG slips

Table 1: Performance of Compression Techniques (Om Prakash, 2012)

Method	CR	CF	SP
RLE	0.384	2.60	61.60
HUFFMAN	0.313	3.19	68.70
LZW	0.224	4.64	77.64
DCT	0.096	10.42	91.68
FFT	0.104	9.62	89.572
DST	0.148	6.76	70.407
DCT-II	0.042	23.81	94.28
FANO	0.684	1.46	31.625
PROPOSED ALGORITHM	0.018	52.8	98.1

The decompression process achieved by two steps first one is to reconstruct the binary image. While the second step focus on projection the data of white color in binary image on the standard ECG paper (paper with baselines). This accomplished by reading the coordinates for each white dot in binary image and draws it (project) as black dot in standard paper at the same coordinate.

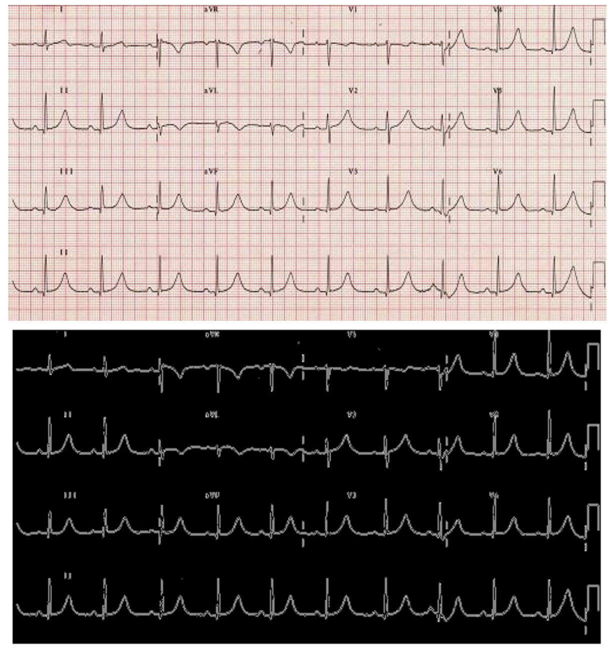
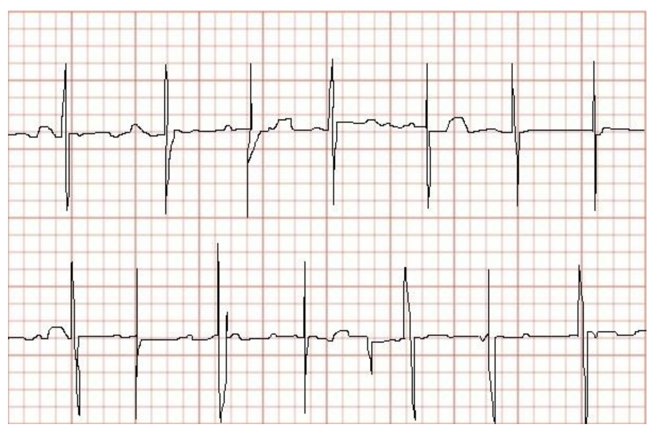
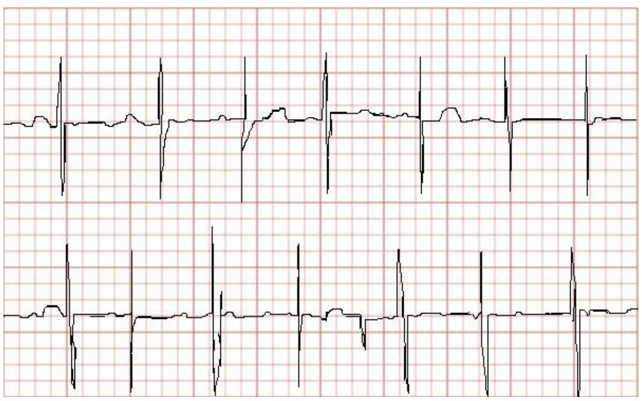


Fig 5: ECG strip and the corresponding binary image



A- Origin image

B- Binary image for image 1



C- Image after decompressed (reconstruct image)

Fig 6: three images before and after compression, and the intermediate step (binary image).

Note: in the Fig 6, Image (A) is drawing image to simulate the ECG strip, and not origin ECG image, just used to measure the performance of algorithm and the quality of the resulted ECG image after decompression..

The quality metrics of proposed compression-decompression algorithm was as in table 2. Where:

PSNR: is the peak signal-to-noise ratio in decibels (dB). The PSNR is only meaningful for data encoded in terms of bits per sample, or bits per pixel.

MSE: The mean square error (MSE) is the squared norm of the difference between the data and the approximation divided by the number of elements.

MAXERR: is the maximum absolute squared deviation of the data (real value signal), from the approximation (reconstructed image).

L2RAT: is the ratio of the squared norm of the signal or image approximation (reconstructed image), to the input signal or image (original image).

Table 2: quality metrics for ECG strip after decompression

Argument	First stage decompression reconstructed binary image	Final stage decompression ECG strip reconstructed
PSNR	∞	63.0108
MSE	0	0.0325
Maxerror	0	1
L2RAT	1	0.6286

There was essentially no false positive diagnosis made on either the compressed or the uncompressed strips, so it can be concluded that the compressor which we evaluated has no measureable influence on diagnostic specificity.

ECG signals that are clean and have a high signal-to-noise ratio (SNR) are relatively easy to interpret, both by a computer and a human healthcare provider.

The new algorithm introduces promise result as highly compression ratio and almost without loss of information visually as fig 6 confirmed. Also, table 2 confirms the similarity of images before and after compression.

With some improvement to this method we can introduce new lossless compression method with high compression ratio.

Digital ECG recording and ECG strip offers potentially higher quality than can be obtained from Holter tape recording, since this method is not subject to wow, flutter, and poor signal-to-noise ratio and low frequency response.

References

- ALSHAMALI, 2010, "Wavelet based ECG compression with adaptive thresholding and efficient coding", Journal of Medical Engineering & Technology, Vol. 34, Nos. 5–6, 335–339.

- ALSHAMALI, M. AL-AQIL, 2011, "ECG compression using wavelet transform and particle swarm optimization", Journal of Medical Engineering & Technology, Vol. 35, No. 3–4, 149–153.

- 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". Digital Signal Processing, 13, 604–622.

- Anubhuti Khare, Manish Saxena, Vijay B. Nerkar, 2011, "ECG Data Compression Using DWT", International Journal of Engineering and Advanced Technology (IJEAT), Volume-1, Issue-1.

- Bendifallah, R. Benzid and M. Boulemden, 2011, "Improved ECG compression method using discrete cosine transform", 3191ELECTRONICS LETTERS Vol. 47 No. 2. DOI 10.1049/el.2010.

- Cardenas-Barrera, J.L. and Lorenzo-Ginori, J.V., 1999, "Mean-shape vector quantizer for ECG signal compression", IEEE Transactions on Biomedical Engineering, 46, 62 – 70.

- Daubechies, I., 1988, "Orthonormal bases of compactly supported wavelets", Communications on Pure & Applied Mathematics, 41, 909–996.

- M. P. S. Chawla, 2009, "A comparative analysis of principal component and independent component techniques for electrocardiograms", Neural Comput & Applic, 18:539–556. DOI 10.1007/s00521-008-0195-1.

- Morteza Moazami-Goudarzi, Mohammad H. Moradi, Ali Taheri, 2005, "Efficient Method for ECG Compression Using Two Dimensional Multiwavelet Transform", World Academy of Science, Engineering and Technology 2.

- Nave G., and Cohen, A., 1993, "ECG compression using long-term Prediction", IEEE Transactions on Biomedical Engineering, 40, 877 – 885.

- Om Prakash, Vivek Chandra, Pushpendra Singh, 2012, "Design and Analysis of an efficient Technique for Compression of ECG Signal", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-1, Issue-5.

- S. M. AHMED, A. F. AL-AJLOUNI, M. ABO-ZAHHAD, and B. HARB, 2009, "ECG signal compression using combined modified discrete cosine and discrete wavelet transforms", Journal of Medical Engineering & Technology, Vol. 33, No. 1, 1–8.

- S. M. AHMED, Q. AL-ZOUBI, and M. ABO-ZAHHAD, 2007, "A hybrid ECG compression algorithm based on singular value decomposition and discrete wavelet transform", Journal of Medical Engineering & Technology, Vol. 31, No. 1, 54 – 61.
- S. M. S. Jalaleddine, C. G. Hutchens, R. D. Strattan and W. A. Coberly. 1990, "ECG Data Compression Techniques – A Unified Approach", IEEE Trans. on Biomedical Eng., vol. 37, 4, 329-341.

Nidhal El Abbadi, received BSc in Chemical Engineering, MSc, and PhD in computer science, worked in industry and many universities, he is general secretary of colleges of computing and informatics society in Iraq, Member of Editorial board of Journal of Computing and Applications, reviewer for a number of international journals, has many published papers and three published books (Programming with Pascal, C++ from beginning to OOP, Data structures in simple language), his research interests are in image processing, biomedical, and steganography, He's Associate Professor in Computer Science in the University of Kufa – Najaf, IRAQ.

Abbas M. Al-Bakry, Graduate from Computer Science Dept., University of Technology-Baghdad in 1989. Get phD. In computer science at 2003, from august 2005 to february 2010 become head of computer science department ,and in 2010 till now work as associate dean of the Information Technology College. Editor in chief topic for the International Journal of Network Computing and Advanced Information Management.