Encyclopedia of Healthcare Information Systems

Nilmini Wickramasinghe
Illinois Institute of Technology, USA

Eliezer Geisler
Illinois Institute of Technology, USA

Volume I
A–D

Medical Information Science Reference

MEDICAL INFORMATION SCIENCE REFERENCE
Hershey • New York

Acquisitions Editor: Kristin Klinger
Development Editor: Kristin Roth
Senior Managing Editor: Jennifer Neidig
Managing Editor: Jamie Snavely
Assistant Managing Editor: Carole Coulson
Copy Editor: Laura Kochanowski, Jeannie Porter, and Sue Vander Hook
Typesetter: Jeff Ash and Sean Woznicki
Cover Design: Lisa Tosheff
Printed at: Yurchak Printing Inc.

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com/reference

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: http://www.eurospanbookstore.com

Library of Congress Cataloging-in-Publication Data

Encyclopedia of healthcare information systems / Nilmini Wickramasinghe and Eliezer Geisler, editors.
p. ; cm.
Includes bibliographical references.
Summary: "This book provides an extensive and rich compilation of international research, discussing the use, adoption, design, and diffusion of information communication technologies (ICTs) in healthcare, including the role of ICTs in the future of healthcare delivery; access, quality, and value of healthcare; nature and evaluation of medical technologies; ethics and social implications; and medical information management"--Provided by publisher.
ISBN 978-1-59904-889-5 (h/c)
1. Medical informatics--Encyclopedias. I. Wickramasinghe, Nilmini. II. Geisler, Eliezer, 1942-
[DNLM: 1. Information Systems--Encyclopedias--English. W 13 E5523 2008]
R858.E52 2008
651.5'04261--dc22

2007047456

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Biomedical Signal Compression

Pedro de A. Berger
University of Brasilia, Brazil

Francisco A. de O. Nascimento
University of Brasilia, Brazil

Leonardo R. A. X. de Menezes
University of Brasilia, Brazil

Adson F. da Rocha
University of Brasilia, Brazil

Joao L. A. Carvalho
University of Southern California, USA

INTRODUCTION

Digitization of biomedical signals has been used in several areas. Some of these include ambulatory monitoring, phone line transmission, database storage, and several other applications in health and biomedical engineering. These applications have helped in diagnostics, patient care, and remote treatment. One example is the digital transmission of ECG signals, from the patient's house or ambulance to the hospital. This has been proven useful in cardiac diagnoses.

Biomedical signals need to be digitally stored or transmitted with a large number of samples per second, and with a great number of bits per sample, in order to assure the required fidelity of the waveform for visual inspection. Therefore, the use of signal compression techniques is fundamental for cost reduction and technical feasibility of storage and transmission of biomedical signals.

The purpose of any signal compression technique is the reduction of the amount of bits used to represent a signal. This must be accomplished while preserving the morphological characteristics of the waveform. In theory, signal compression is the process where the redundant information contained in the signal is detected and eliminated. Shannon (1948) defined redundancy as "the fraction of unnecessary information, and therefore repetitive in the sense that if it was missing, then the information would still be essentially complete, or it could at least be recovered."

Signal compression has been widely studied during the past decades, and several references discuss this subject (Gersho & Gray, 1992; Jayant & Noll, 1982; Sayood, 1996).

Signal compression techniques are commonly classified in two categories: lossless and lossy compression. Lossless compression means that the decoded signal is identical to the original one. In lossy compression, a controlled amount of distortion is allowed. Lossy signal compression techniques show higher compression gains than lossless ones.

BACKGROUND

Lossless Compression

Lossless signal compression techniques are less efficient with respect to compression gains. They may be used in combination with lossy compression techniques, especially in cases where the maximum allowed distortion has been reached, and additional compression is needed. Among several lossless compression techniques, we highlight entropy coding (Gersho & Gray, 1992), Run-Length, Huffman (Huffman, 1952), arithmetic coding (Witten, 1987), and delta coding (Ken, 1985).

Run-Length Coding

Data files frequently present sequentially repeated characters (a character run). For instance: text files use several spaces to separate sentences and paragraphs. Digital signals may contain the same value, or the same character representing that value in its data file, repeated many times sequentially. This indicates that the signal is not changing, as in the isoelectric segments of ECG signals, for example.

Figure 1 shows an example of Run-Length coding of a data set that contains runs of zeros. Each time the coder finds a zero in the entry data, two values are written in the output data. The first of these values is a zero indicating that the Run-Length codification started. The second value is the amount of zeros in the sequence. If the run of zeros in the input data set is in average larger than two, then the Run-Length coder will achieve data compression.

Huffman Coding

In Huffman coding, the data are represented as a set of variable length binary words. The lengths depend on the occurrence frequency of the symbols used for representing each signal value. Characters that are used often are represented with fewer bits, and those that are seldom used are represented with more bits.

Figure 2 shows an example of how a Huffman code is generated, given a data set X and its characteristic probability of occurrence—p(X). The character codes are generated by combining the bits of a tree with ramifications, adding their probabilities and restarting the process until only one character remains. This process generates a three with ramifications linked to bits 0 and 1. The codes for each character are taken in the inverse path of these ramifications. Note that initial character arrangement is not relevant. In this example, we chose to encode the upper ramifications with bit 0 and the lower ones with bit 1. However, the opposite representation could have been used as well. Any decision criteria may be used in ramifications with equal probabilities.

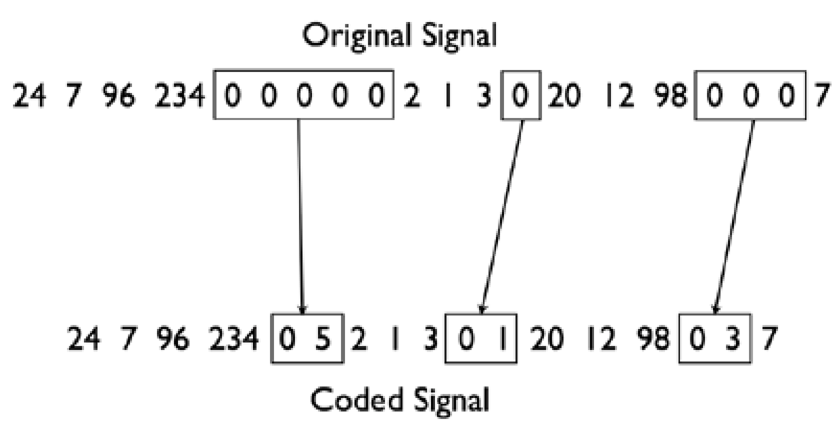
Huffman coding has the disadvantage of assigning an integer number of bits to each symbol. This is a suboptimal strategy, because the optimal number of bits per symbol depends on the information content, and is generally a rational number. Arithmetic coding is a more sophisticated compression technique, based on Huffman coding concepts. It results in compression gains closer to the theoretical limits. In this coding, character sequences are represented by individual codes depending on their probability of occurrence. Huffman and Arithmetic codes are often used in combination with Run-Length coding, for further compressing biomedical signals that were first compressed using orthogonal transforms.

Delta Coding

Delta coding refers to signal compression techniques that store a digital signal as the difference between successive samples. Figure 3 shows an example of how this is performed. The first sample in the encoded signal is equal to the first sample of the original signal. Subsequent encoded samples are equal to the difference between the current sample and the previous sample of the original signal.

Using this technique, the encoded signal has a smaller amplitude dynamic range than the original signal. Therefore, it takes fewer bits to store or transmit the encoded signal. Delta coding is a particular case of Differential Pulse Code Modulation (DPCM). DPCM is used in combination with Huffman coders in several biomedical signal compression algorithms.

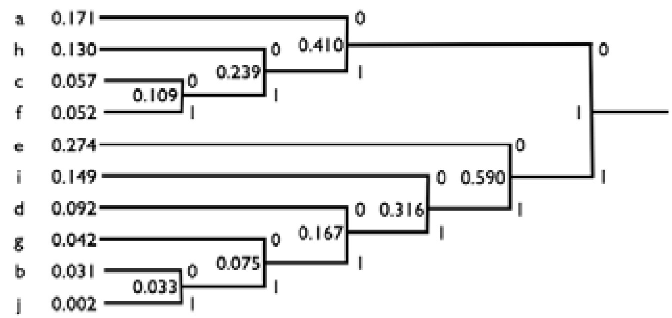
Figure 1. Run-length coding example



B

Figure 2. Huffman code construction example

X	p(X)
a	0.171
b	0.031
c	0.057
d	0.092
e	0.274
f	0.052
g	0.042
h	0.130
i	0.149
j	0.002



X	Code
a	00
b	111110
c	0110
d	1110
e	10
f	0111
g	11110
h	010
i	1100
j	111111

Figure 3. Delta coding example

Original Signal: 23 20 19 21 11 75 79 84 ...

Coded Signal: 23 -3 -1 2 -10 64 4 5 ...

Lossy Compression

There are two major categories of lossy compression techniques used with biomedical signals: direct methods, and transform compression.

Direct Methods

Direct methods encode signals in time domain. These algorithms depend on the morphology of the input signal. In most cases, these methods are complex, and provide lower compression efficiency than transform coders (discussed in the next section).

Direct compression methods use sophisticated processes and "intelligent" signal decimation and interpolation. In other words, these methods extract K "significant" samples of the original signal x(n), such that:

$$(n, x(n)),\ n = 0,\ldots,N-1 \rightarrow (n_k, x(n_k)),\ k = 0,\ldots,K-1, \tag{1}$$

where k<N. The reconstruction of values between significant samples is performed by interpolation, using the generic expression (Sörnmo & Laguna, 2006):

$$\hat{x}(n) = \begin{cases} x(n), & n = n_0,\ldots,n_{k-1}; \\ f_{n_0,n_1}(n), & n = n_0+1,\ldots,n_1-1; \\ \vdots & \vdots \\ f_{n_{K-2},n_{K-1}}(n) & n = n_{K-2}+1,\ldots,n_{K-1}-1; \end{cases} \tag{2}$$

The process of selecting significant samples is based on the characteristics of the signal, and on a tolerance interval criterion for the reconstruction error. The interpolation function $f_{n_{k-1},n_k}(n)$ is generally a low order polynomial, of either zero or first order. These polynomials approximate the original signal with a series of lines connecting the K significant samples. The geometric parameters of these lines (length and inclination) are stored instead of the discarded sample values, resulting in a reduction in the amount of stored or transmitted data.

Transform Compression

Among the several methods for signal compression, the techniques based on transforms yield the best performance in terms of compression gain and fidelity of the waveform of the reconstructed signal.

Given a vector of data $x=[x(1)\ x(2) \ldots.\ x(N)]$, we can define an orthogonal transform as a linear operation given by a linear transformation T, such that:

$$y = Tx, \tag{3}$$

where $y=[y(1)\ y(2)\ \ldots\ y(N)]$ represents a vector of transformed coefficients, and T satisfies the orthogonality condition:

$$T^t = T^{-1} \tag{4}$$

Transform compression is based on a simple premise: when the signal is processed by a transform, the signal energy (information) that was distributed among all the samples in the time domain can be represented with a small number of transformed coefficients.

This can be observed in Figure 4, where an ECG signal is shown along with the correspondent coefficients in transformed domain. In this example, we use the discrete cosine transform (DCT). The DCT is used in the most popular standard for image compression, the *Joint Picture Expert Group* (JPEG).

Note that, in the transformed domain, the energy of the signal is concentrated in a small number of coefficients with high amplitude. Thus, if we store the high-amplitude coefficients, and discard the low-amplitude coefficients, the signal may be represented accurately enough, and can be recovered by inverse transformation. Figure 5 shows the same ECG signal in Figure 4, and the reconstructed signal after discarding 80% of the DCT coefficients. Note that, by storing only 20% of the DCT coefficients it is possible to reconstruct the original signal without visible distortions.

Currently, the most widely used transform for encoding biomedical signals is the discrete wavelet transform (Daubechies, 1988; Mallat, 1989). The wavelet transform is used for image compression—such as in the JPEG 2000 compression standard—and in more recent studies on biomedical signal compression. In this transform, a signal containing N samples is filtered by a pair of filters that decompose the signal into low- (L) and a high- (H) frequency bands. Each band is undersampled by a factor of two; that is, each band contains N/2 samples. With the appropriate filter design, this action is reversible. This procedure can be extended for two-dimensional signals, such as images. In Figure 6, we show an example of wavelet decomposition for a gray-scale image with 256 x 256 pixels. Similarly to what was observed using the DCT, many of the coefficients in the high-frequency subbands have amplitudes close to zero (very dark pixels), and it is possible to compress the image by discarding them.

Wavelet transform compression is evolving with respect to the way in which the coefficients are encoded. When a coefficient in a low-frequency subband is nonzero, there is a high probability that, at positions that correspond to high frequencies, the coefficients are also nonzero. Thus, the nonzero coefficients can be represented in a tree, beginning with a low frequency root. Figure 7 illustrates this concept. A single coefficient in the LL band of layer 1 has a correspondent coefficient in the other bands. The positions of the coefficients in layer 1 are mapped into four daughter-positions in each subband of layer 2.

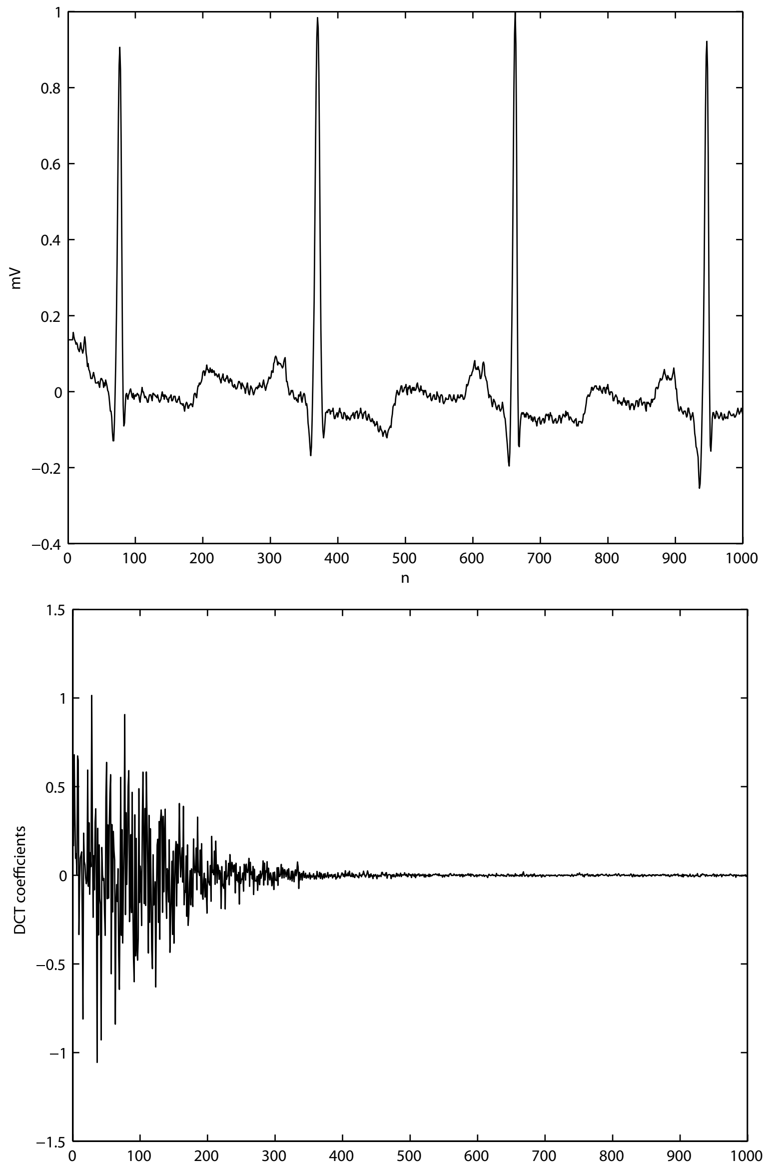
An efficient way of coding the coefficients that are non-zero is to code each tree of coefficients, beginning with the root decomposition level. The coefficients at the lower layers are encoded, and followed by their children coefficients in the higher layer, until a null coefficient is found. The next coefficients of the tree have great probability of also being null, and they are replaced by a code that identifies a tree of zeros (zerotree code). This method is called Embedded Zerotree Wavelet (EZW). Another method, similar to the EZW, and commonly used in the wavelet coefficients encoding, is the Set Partitioning In Hierarchical Tree (SPIHT) (Said & Pearlman, 1996).

COMMENTS ON CURRENT AND FUTURE RESEARCH

The compression of biomedical signals has been extensively studied by the scientific community. In this article, we discuss more thoroughly the compression of electrocardiographic (ECG) and electromyographic (EMG) signals.

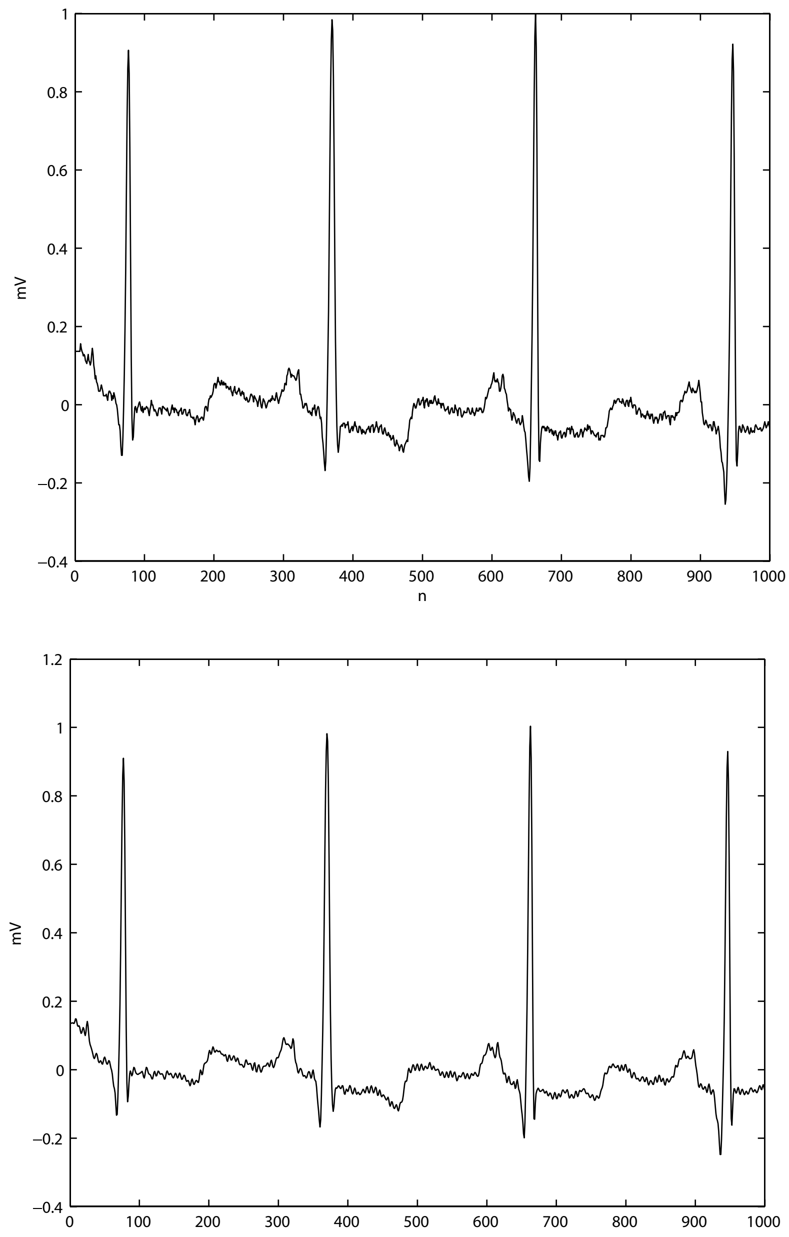
In general, we can group the techniques for compression of ECG signals in two main groups:

Figure 4. (a) ECG signal (top); (b) transformed coefficients (bottom)



1. Dedicated techniques:
   a. Direct Methods: Amplitude Zone Time Epoch Coding (AZTEC) (Cox, Noelle, Fozzard, & Oliver, 1968), Turning Point (TP) (Mueller, 1978), Coordinate Reduction Time Encoding System (Abenstein & Tompkins, 1982), FAN algorithm (Dipersio & Barr, 1985) and modified time domain coding algorithms, such as SLOPE (Tai, 1991) and Scan Along Polygonal Approximation (SAPA) (Shahein & Abbas, 1994).
   b. Optimization algorithms: Long-Term Prediction (LTP) (Nave & Cohen, 1993), algorithms based on analysis by synthesis (ASEC) (Jalaleddine, 1990), and the Cardinality Constrained technique (Nygaard & Hauglan, 1998).
2. Generic techniques: These techniques can be used with a great variety of signals, including voice, image, and video. This includes Differential Pulse Code Modulation (DPCM), Subband Coding (SC) and transform-based coding (Hilton, 1997; Istepanian, Hadjileontiadis, & Panas, 2002; Lu, Kim, & Pearlman, 2000; Rajoub, 2002).

Figure 5. (a) original ECG signal (top); (b) reconstructed ECG signal after discarding 80% of the coefficients (bottom)



Currently, the algorithms that provide the best ECG compression results use transform-based coding, and the most widely used is the wavelet transform. These algorithms can be grouped, with respect to coefficient encoding, into two main groups: algorithms with coefficient thresholding, and algorithms with tree-based coefficient encoding.

Coefficient thresholding algorithms follow a methodology that was presented in the previous section. The majority of these algorithms also apply lossless compression techniques to encode the wavelet coefficients. One example of this technique is the work by Rajoub (2002).

Tree-based coefficient encoding algorithms use hierarchic structures for encoding the wavelet coefficients, by exploiting the redundancy in the scales of the wavelet decomposition. Such algorithms provide results superior to traditional thresholding algorithms. Among recent works that use this paradigm, we can mention the works of Istepanian et al. (2002), Hilton (1997), and Lu et al. (2000).

B

Figure 6. Two-dimensional wavelet decomposition

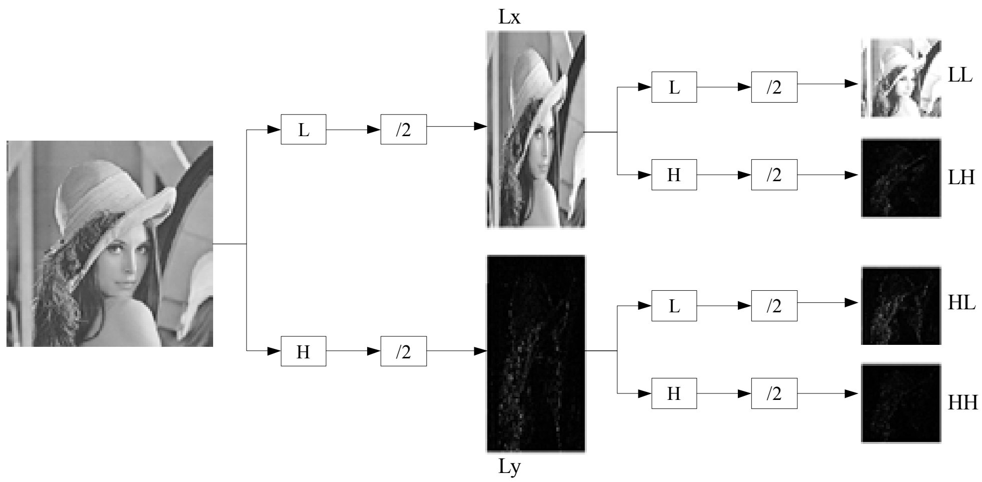
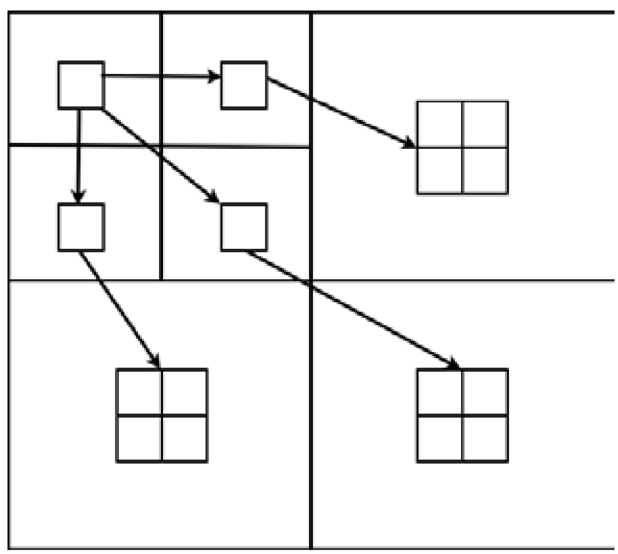


Figure 7. Wavelet coefficients tree



The main techniques described in the literature for EMG signal compression can be found in (Berger, Nascimento, Carmo, & Rocha, 2006; Guerrero & Maihes, 1997; Norris, 2001; Norris, Englehart, & Lovely, 2001; Wellig, Zhenlan, Semling, & Moschytz, 1998). Norris et al. (1995) investigates the compression of EMG signals using Adaptive Differential Pulse Code Modulation (ADPCM). Guerrero and Maihes (1997) compared different methods for lossless compression with methods based on orthogonal transforms. The orthogonal transform based showed better performance, with respect to compression ratio.

Recently, the use of wavelet transform for compression of EMG signals was studied by Wellig et al. (1998), Norris et al. (2001), and Berger et al. (2006). Wellig et al. (1998) and Norris et al. (2001) encoded the transformed coefficient using the EZW algorithm and compared their results with the traditional wavelet transform. Berger et al. (2006) used a scheme for adaptive dynamic bit allocation for encoding the coefficients. These methods provide performance equivalent or better than previously described methods.

Comparing different algorithms for EMG signal compression is difficult, because there is no standard database, such as the *MIT-BIH Arrhythmia Database*, which is used in most studies on compression of ECG signals. The studies on EMG signal compression usually use different test signals, involving different muscles, volume conductors, conduction velocities, loads, sampling rates and duration. Therefore, the comparison between different algorithms should always be analyzed with caution.

The algorithms mentioned in this article have the goal of guaranteeing the fidelity of the reconstructed signal after decoding. In this sense, the wavelet transform and the EZW and SPIHT algorithms are considered by the specialized public as the most promising lines of work, and future research tend to emphasize improvements in these algorithms.

However, reaching a good quality representation of signals with low bit rates comes with the cost of increasing the complexity of the compression algorithms. Few studies on low-cost real-time compression of biomedical signals have been published, even though the interest on telemedicine applications is growing. This is an important issue, which should be addressed in the future. The works by Alesanco, Olmos, Istepanian, and García (2003) and Kim, Yoo, and Lee (2006) proposed the first solutions for real-time ECG compression, and Carotti, De Martin, Merletti, and Farina (2006) a lossy coding technique for surface EMG signals that is based on the algebraic code excited linear prediction (ACELP) paradigm, widely used for speech signal coding. The algorithm was adapted to the EMG characteristics and tested on both simulated and experimental signals. This method is characterized by moderate complexity (approximately 20 million instructions/s) and an algorithmic delay smaller than 160 samples (~160 ms).

CONCLUSION

This article presented a review of techniques for compression of biomedical signals. Among the different techniques that were presented, the ones based on wavelet transforms had the best performance with respect to compression gain and reconstructed waveform fidelity. When the wavelet transform is used, a great challenge is to find effective alternatives for efficiently encoding the transformed coefficients. The most efficient methods for this are the EZW and SPIHT algorithms. An important line of future work is the search for real-time, low-consumption algorithms, which would be suitable for telemedicine applications.

REFERENCES

Abenstein, J., & Tompkins, W. (1982). New data-reduction algorithm for real-time ECG analysis. *IEEE Trans. Biomed. Engineering. 29*, 43–48.

Alesanco, A., Olmos, S., Istepanian, R., & Garcia, J. (2003). A novel real-time nultilead ECG compression and de-noising method based on the wavelet transform. *Computers in Cardiology*, 30, 593-596.

Alesanco, A., Olmos, S., Istepanian, R.S.H., & García, J. (2006) Enhanced real-time ECG coder for packetized telecardiology applications. *IEEE Transactions on Information Technology in Biomedicine, 10*(2), 229–236.

Berger, P.A., Nascimento, F.A.O., Carmo, J.C., & Rocha, A.F. (2006) Compression of EMG signals with wavelet transform and artificial neural networks. *Physiological Measurement, 27*, 457–465.

Carotti, E., De Martin, J.C., Merletti, R., & Farina, D. (2006). Compression of surface EMG signals with algebraic code excited linear prediction. *Medical Engineering & Physic, 29* (2), 253–258.

Cox, J., Noelle, F., Fozzard, H., & Oliver, G. (1968). AZTEC: A preprocessing program for real-time ECG rhythm analysis. *IEEE Trans. Biomed. Engineering. 15*, 128–129.

Daubechies, I. (1988). Orthogonal bases of compactly supported wavelets. *Comm. Pur. Appl. Math. 41*, 909–996.

Dipersio, D.A., & Barr, R.C. (1985). Evaluation of the FAN method of adaptive sampling on human eletrocardiograms. *Med. Biomed. Eng. Computing*, 401–410.

Guerrero, A., & Maihes, C. (1997). On the choice of an electromyogram data compression method. In *Proceedings of the 19th annual International Conference of the IEEE Engineering in Medicine Biology Society* (pp. 1558–1561).

Gersho, A., & Gray, R. (1992). *Vector quantization and signal compression*. Kluwer Acad. Publication.

Hilton, M.L. (1997). Wavelet and wavelet packet compression of electrocardiograms. *IEEE Trans. Biomed. Engineering, 44*, 394–402.

Huffman, D.A. (1952). A method for the construction of minimum redundancy codes. In *Proceedings of IRE* (Vol. 40, No. 10, pp. 1098–1101).

Istepanian, R.S.H., Hadjileontiadis, L., & Panas, S. (2001). ECG data compression using wavelets and higher order statistics methods for telemedical applications. *IEEE Transactions on Information Technology in Biomedicine, 5*(2), 108–115.

Jalaleddine, S.M., Hutchens, C.G., Satrattan, R.D., & Coberly, W.A. (1990). ECG data compression techniques—The unified approach. *IEEE Trans. Biomed. Engineering*, *37*(4), 329-343.

Jayant, N.S., & Noll, P. (1982). *Digital coding of waveforms – principles and application to speech and video*. Prentice Hall.

Ken, C. P. (1985). *Principles of digital audio*. Sams/Prentice-Hall Computer Publishing.

Kim, S.B., Yoo, S.K., & Lee, M.H. (2006). Wavelet-based low-delay ECG compression algorithm for continuous ECG transmission. *IEEE Transactions on Information Technology in Biomedicine, 10*(1), 77–83.

Lu, Z., Kim, D.Y., & Pearlman, W.A. (2000). Wavelet compression of ECG signals by the set partitioning in hierarchical trees (SPIHT) algorithm. *IEEE Trans. Signal Processing, 47*, 849–856.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. In *Proceedings of IEEE Trans on Pattern Anal. and Math. intel.* (Vol. 11, No. 7).

Mueller, W. (1978). Arrhythmia detection program for an ambulatory ECG monitor. *IEEE Trans. Biomed. Engineering, 14*, 81-85.

Nave, G., & Cohen, A. (1993). ECG compression using long-term prediction. *IEEE Trans. Biomed. Engineering, 40*, 877-885.

Norris, J.A., Englehart, K., & Lovely, D. (2001). Steady-state and dynamic myoelectric signal compression using embedded zero-tree wavelets. In *Proceedings of the 23th annual International Conference of the IEEE Engineering in Medicine Biology Society* (pp. 1879–1882).

Nygaard, R., & Haugland, D. (1998). Compressing ECG signals by piecewise polynomial approximation. In *Proceedings of the Int. Conf. Acoust. Speech Signal Processing* (pp. 420-424). Seattle, WA.

Rajoub, B.A. (2002). An efficient coding algorithm for the compression of ECG signals using the wavelet transform. *IEEE Trans. Biomed. Eng., 49*(4), 355–362.

Said, A., & Pearlman, W.A. (1996). A new fast and efficient image codec based on set partitioning in hierarchical Trees. *IEEE Transactions on Circuits and Systems for Video Technology, 6*, 243–250.

Sayood, K. (1996). Introduction to data compression. In C.F. Shahein, & H.M. Abbas (Eds.), ECG Data-Compression via Cubic-Splines and Scan Along Polygonal Approximation. *Signal Processing, 35*(3), 269–283. Morgan Kaufman.

Shahein, C.F., & Abbas, H.M. (1994). ECG data-compression via cubic-splines and scan along polygonal approximation. *Signal Processing, 35*(3), 269-283.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Tech. Journal, 27*, 379-423.

Sörnmo, L., & Laguna, P. (2006). Electrocardiogram (ECG) signal processing. In M. Akay (Ed.), *Wiley Encyclopedia of Biomedical Engineering* (Vol. 2, pp. 1298–1313). Wiley.

Tai, S.C. (1991). Slope – a real-time ECG data compressor. *L Biomed. Eng. Computing, 29*, 175–179.

Wellig, P., Zhenlan, C., Semling, M., & Moschytz, G.S. (1998). Electromyogram data compression using single-tree and modified zero-tree wavelet encoding. In *Proceedings of Engineering in Medicine and Biology Society 23rd Annual International Conference of the IEEE* (pp. 1303–1306).

Witten, I., Neal, R., & Cleary, J. (1987). Arithmetic coding for data compression. *Communications of the ACM, 30*(6).

KEY TERMS

Arithmetic Coding: A more sophisticated compression technique, based on Huffman coding concepts. It results in compression gains closer to the theoretical limits. In this coding, character sequences are represented by individual codes depending on their probability of occurrence.

Biomedical Signals: Refers to signals that carry useful information for probing, exploring, and understanding the behavior of biomedical systems under investigation.

Delta Coding: Refers to signal compression techniques that store a digital signal as the difference between successive samples. The first sample in the encoded signal is equal to the first sample of the original signal. Subsequent encoded samples are equal to the difference between the current sample and the previous sample of the original signal.

Direct Methods: Encode signals in time domain. These algorithms depend on the morphology of the input signal. Direct compression methods use sophisticated processes and "intelligent" signal decimation and interpolation.

Discrete Cosine Transform: A technique for expressing a waveform as a weighted sum of cosines. The DCT is central to many kinds of signal processing, especially video compression.

Embedded Zerotree Wavelet: A simple, yet remarkable effective, compression algorithm witch exploits the self-similarity of the wavelet transform coefficients at different scales, by structuring the data in hierarchical tree sets that are likely to be highly correlated, yielding a fully embedded code.

Huffman Coding: Refers to a compression technique where the data are represented as a set of variable length binary words. The lengths depend on the frequency of occurrence of the symbols used for representing each signal value. Characters that are used often are represented with fewer bits, and those that are seldom used are represented with more bits.

Lossless Compression: A signal compression technique where the original signal can be perfectly recovered from the compressed signal. The Morse code of telegraphy provides a classic example: by sending short patterns (dots and dashes) for frequent letters, and long patterns for rare letters, fewer bits are needed on the average than with the standard computer representation (ASCII), which assigns the same number of bits to all letters.

Lossy Compression: Compression is lossy if the original signal cannot be perfectly recovered from the compressed file, as a controlled amount of distortion is allowed. Lossy signal compression techniques show higher compression gains than lossless ones.

Run-Length Coding: A lossless compression. A set of symbols is represented as a sequence of RUN/LEVEL. RUN being the number of consecutive zeros and LEVEL being the value of the following nonzero coefficient.

Set Partitioning in Hierarchical Tree: A refinement of the Embedded Zerotree Wavelet algorithm. SPIHT requires partitioning the wavelet coefficients into a number of lists, with list membership changing as the execution proceeds, and in some cases involving dual membership of a coefficient in different lists.

Signal Compression: The reduction of the necessary amount of bits to represent a signal. In many cases, this must be accomplished while preserving the morphological characteristics of the waveform. In theory, signal compression is the process where the redundant information contained in the signal is detected and discarded.

Transform Compression: Compression technique that is based on a simple premise: when the signal is processed by a transform, the signal energy (information) that was distributed among all the samples in the time domain can be represented with a small number of transformed coefficients.

Wavelet Transform: Refers to a transformation where a signal containing N samples is first filtered by a pair of filters that decompose the signal into low and a high frequency bands. Each band is undersampled by a factor of two, that is, each band contains N/2 samples. With the appropriate filter design, this action is reversible.