

Compresión de fonocardiogramas mediante las transformadas wavelet y wavelet packet

Juan Martínez Alajarín, Ramón Ruiz Merino

Departamento de Electrónica, Tecnología de Computadoras y Proyectos

Universidad Politécnica de Cartagena, ETSII, Campus Muralla del Mar, C/ Dr. Fleming s/n, 30202 Cartagena

Tlf.: 968 32 64 64, 968 32 64 53. Fax: 968 32 64 00. E-mail: juanc.martinez@upct.es, ramon.ruiz@upct.es

Resumen

En este artículo se describe un método específico para la compresión del fonocardiograma. Este método emplea la transformada wavelet, descartando aquellos coeficientes inferiores a un determinado umbral (compresión con pérdidas) junto a otros métodos adicionales para aumentar la tasa de compresión, como la codificación Huffman y la codificación RLE (compresión sin pérdidas). Los resultados muestran mejor tasa de compresión y menor distorsión que las obtenidas mediante las técnicas estándar de compresión de audio, debido a que el método propuesto está específicamente diseñado para la banda de frecuencias de interés de las señales fonocardiográficas.

1. Introducción

Tras algunas décadas en las que la auscultación del corazón perdió importancia a favor de otras técnicas más precisas (y también costosas), en los últimos años parece que se ha redescubierto la utilidad del fonocardiograma (FCG) como una técnica rápida, sencilla y de bajo coste de ayuda al diagnóstico de enfermedades valvulares [1]. El desarrollo de la Telemedicina puede ser muy útil para la auscultación automatizada de los sonidos cardíacos con el fin de transmitir la señal FCG para que sea analizada y procesada remotamente.

Hasta la fecha, casi todos los trabajos realizados en el ámbito de la compresión de señales biomédicas se han centrado en el electrocardiograma (ECG), para el que existen numerosas técnicas específicas de compresión [2], algunas de ellas utilizando las transformadas wavelet (WT) y wavelet packet (WPT) [3-5]. Sin embargo, hasta donde los autores conocen, no existe bibliografía donde se describan trabajos relacionados con la compresión del FCG.

Puesto que esta señal consiste en la representación gráfica de los sonidos y soplos cardíacos generados por el corazón, podría pensarse en un principio en utilizar técnicas estándar de compresión de audio. Sin embargo, estas técnicas están orientadas principalmente a la compresión con gran calidad de música o del habla, por lo que su aplicación a la compresión del FCG no parece adecuada, dado que el rango de frecuencias de interés de la música o del habla es muy diferente al del FCG.

En este artículo se propone un método eficiente para comprimir la señal FCG mediante la eliminación de aquellos coeficientes wavelet, obtenidos de la WT o de la

WPT, inferiores a un determinado umbral (compresión con pérdidas). Con el fin de aumentar aún más la tasa de compresión, también se han empleado técnicas tradicionales de compresión sin pérdidas, como la codificación Huffman y la codificación Run Length Encoding (RLE).

2. Descripción del algoritmo

El algoritmo propuesto para la compresión del FCG consta de las siguientes cuatro etapas (Figura 1):

- Descomposición del FCG mediante la WT o WPT.
- Aplicación de un umbral a los coeficientes wavelet.
- Compresión del vector de los coeficientes wavelet, empleando eliminación de ceros y cuantización lineal.
- Compresión del mapa de posición, empleando eliminación de ceros, y las codificaciones RLE y Huffman.

La descompresión (Figura 2) realiza los pasos inversos a los indicados, con la única diferencia de que no existe el paso inverso a la umbralización, que es el que provoca las pérdidas en la compresión:

- Descompresión del vector de coeficientes.
- Descompresión del mapa de posición.
- Reconstrucción de la señal a partir del vector de coeficientes wavelet.

A continuación se describe con más detalle cada una de las cuatro etapas de la compresión.

2.1. Compresión wavelet

Inicialmente, el FCG se divide en bloques no solapados de L_b muestras, y cada bloque se comprime independientemente de los restantes. Tanto la WT como la WPT descomponen la señal de cada bloque en $M+1$ y 2^M subbandas, respectivamente, donde M es el nivel de profundidad de la descomposición, de donde se extrae un vector de coeficientes para cada subbanda. Para la WPT, el mejor árbol de descomposición [6] se determina mediante el criterio de la entropía de Shannon no normalizada.

Sea \mathbf{WC} el vector formado por los coeficientes de todas las subbandas, ordenadas de mayor a menor escala. Para la WT, el vector \mathbf{WC} se define como:

$$\mathbf{WC}^T = (\mathbf{CA}_M^T, \mathbf{CD}_M^T, \mathbf{CD}_{M-1}^T, \dots, \mathbf{CD}_1^T)$$

donde \mathbf{CA}_M es el vector de los coeficientes de aproximación, y de \mathbf{CD}_M a \mathbf{CD}_1 son los vectores de los coeficientes de detalle.

Entonces se calcula un umbral Thr de forma que retenga un porcentaje determinado de la energía de la señal original ($RtEn$). Tras aplicar este umbral al vector \mathbf{WC} , se obtienen dos vectores. El primero de ellos es el vector de coeficientes umbralizados (\mathbf{TC}), que es el mismo vector que \mathbf{WC} pero con los coeficientes que tienen valores absolutos iguales o menores que Thr hechos nulos. El segundo vector es el mapa de posición (\mathbf{SM}), que consta de un mapa binario que indica con '1's y '0's la posición de los coeficientes no nulos y nulos, respectivamente, en el vector \mathbf{WC} .

De aquí se derivan otros dos vectores: el vector de coeficientes umbralizados no nulos (\mathbf{NZTC}), en el que se han eliminado los coeficientes nulos del vector \mathbf{TC} , y el mapa de posición no nulo (\mathbf{NZSM}), en el que el último bloque de '0's se ha eliminado.

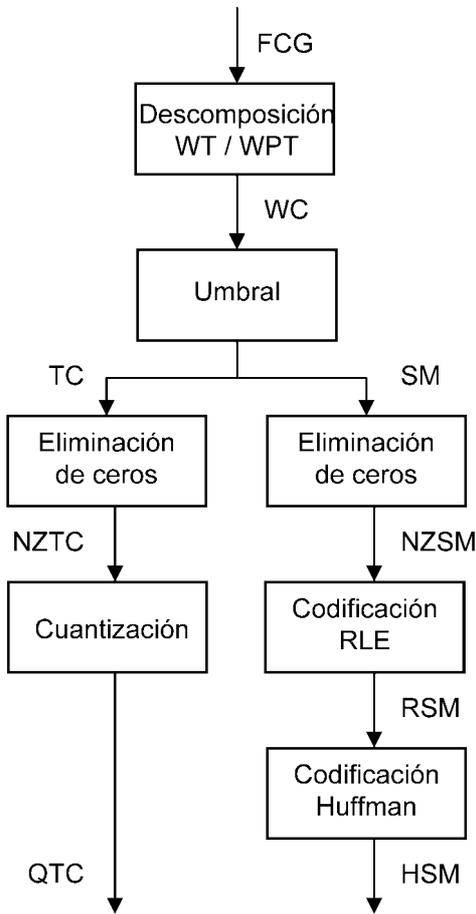


Fig. 1. Diagrama de bloques del algoritmo de compresión del FCG

2.2. Compresión adicional

Con el fin de obtener una mayor tasa de compresión, los vectores \mathbf{NZTC} y \mathbf{NZSM} se han comprimido con técnicas adicionales de forma independiente, dada la diferente naturaleza de ambos vectores.

En primer lugar, el vector \mathbf{NZTC} se cuantiza linealmente en el rango $0-(2^{Nb}-1)$, mediante la expresión:

$$\mathbf{QTC} = 2^{Nb} \cdot \frac{\mathbf{NZTC} - \mathbf{NZTC}_{\min}}{\mathbf{NZTC}_{\max} - \mathbf{NZTC}_{\min}}$$

donde Nb es el número de bits empleados para la cuantización, \mathbf{QTC} es el vector de coeficientes cuantizados, y \mathbf{NZTC}_{\min} y \mathbf{NZTC}_{\max} son, respectivamente, los valores mínimo y máximo de \mathbf{NZTC} .

Por otra parte, el vector \mathbf{NZSM} se codifica primero mediante RLE, proporcionando el vector \mathbf{RSM} , y después con el algoritmo Huffman, resultando el vector \mathbf{HSM} .

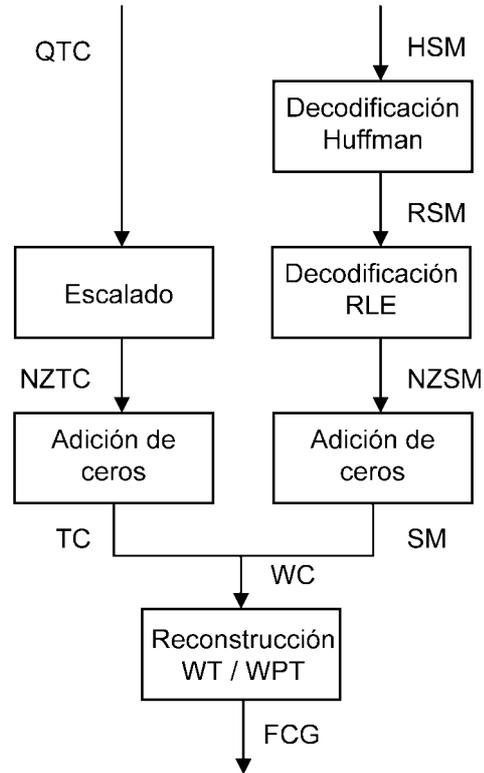


Fig. 2. Diagrama de bloques del algoritmo de descompresión del FCG

2.3. Evaluación de la compresión

Sea \mathbf{X} una señal FCG de N muestras que se va a comprimir. La compresión obtenida respecto a la señal original se evalúa normalmente mediante la tasa de compresión y la distorsión producida al comprimir la señal.

La tasa de compresión (CR) se define como:

$$CR = \frac{L_x}{L_c}$$

donde L_x es el tamaño de la señal original y L_c es el tamaño de la señal comprimida (\mathbf{QTC} y \mathbf{HSM}).

Por otra parte, la distorsión generada por la compresión (o error de reconstrucción) se define como la raíz cuadrada de la diferencia cuadrática media (PRD, Percent Root-mean-square Difference), modificada para ser independiente de la media de \mathbf{X} :

$$PRD = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{\sum_{i=1}^N (x_i - \mathbf{m}_x)^2}} \times 100$$

donde x_i es la señal original, \hat{x}_i es la señal reconstruida, y \mathbf{m}_x es la media de \mathbf{X} .

Finalmente, hay que indicar que la compresión depende de los siguientes parámetros: a) la función madre wavelet (Wmf), b) el nivel (o profundidad) de descomposición WT o WPT (M), c) el número de bits para la cuantización lineal (Nb), d) el número de muestras de cada bloque (Lb), y e) el porcentaje de energía retenida ($RtEn$).

3. Resultados experimentales

Para comprobar las prestaciones de este método de compresión se ha aplicado a 52 registros FCG reales almacenados en formato WAV (16 bits, mono y 8000Hz). Esta base de datos incluye registros de sonidos normales y de diversas patologías valvulares (tercer ruido, cuarto ruido, clics, soplos sistólicos, diastólicos y continuos, etc.).

En un primer paso se realizó la compresión con distintas funciones wavelet, obteniendo que la que conseguía una mayor tasa de compresión con menor error y tiempo de computación era la función $Wmf = Daubechies8$. Posteriormente se realizaron pruebas combinando distintos valores para el resto de parámetros que intervienen en la compresión, obteniendo los mejores resultados para los valores $M = 4$, $Lb = 4096$, y $Nb = 10$. Las Figuras 3 y 4 muestran, respectivamente, los valores de CR y PRD para los registros de prueba para este conjunto de parámetros, para $RtEn = 99.99\%$.

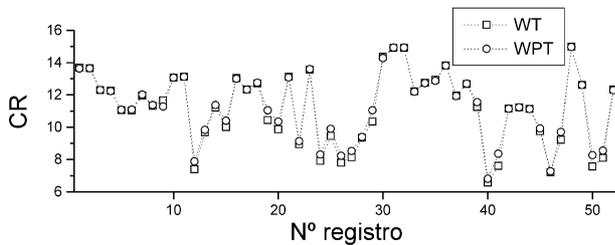


Fig. 3. Valores de CR para 52 registros FCG comprimidos con WT y WPT, para $RtEn = 99.99\%$

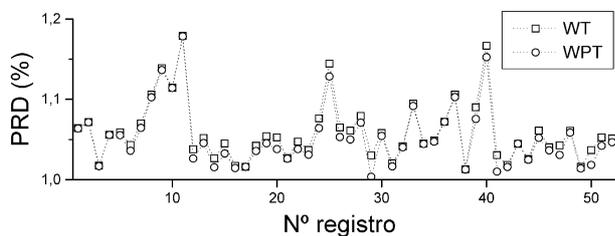


Fig. 4. Valores de PRD para 52 registros FCG comprimidos con WT y WPT, para $RtEn = 99.99\%$

La Tabla 1 muestra la media y la desviación estándar para los valores de CR y PRD y la relación entre el tiempo de compresión de la WPT y de la WT, promediados para los 52 registros, obtenidos para $RtEn = 99.9\%$, 99.99% , y 99.999% , con el resto de parámetros fijos, para la WT y la WPT. Esta tabla muestra que los resultados obtenidos con

la WT y la WPT son muy similares en términos de CR y PRD, mientras que el tiempo de cálculo de la WPT es considerablemente superior al de la WT. Por esta razón, y por la mayor simplicidad en el cálculo de la WT respecto a la WPT (no hay que calcular el mejor árbol de descomposición), se ha elegido en nuestro caso la WT como método de cálculo preferido.

$RtEn$ (%)	WT		WPT		$\frac{T_{WPT}}{T_{WT}}$
	CR	PRD	CR	PRD	
99.900	16.75± 3.17	3.156± 0.078	16.71± 3.09	3.147± 0.078	1.76
99.990	11.19± 2.23	1.059± 0.038	11.33± 2.07	1.053± 0.038	1.55
99.999	7.01± 1.62	0.542± 0.084	7.12± 1.61	0.535± 0.081	1.35

Tabla 1. Media y desviación típica de los valores de CR y PRD (%) para WT y WPT, y relación del tiempo de compresión entre WPT y WT, promediados para 52 registros FCG

Finalmente se ha realizado una comparación entre el método propuesto y el método de compresión de audio OGG Vorbis. Para ello se ha comprimido con este último método un registro FCG correspondiente a un sonido normal, de duración 13.9s, variando su calidad desde -0.1 hasta 1 en incrementos de 0.1 (Tabla 2). Los valores de CR se han calculado como el cociente entre el tamaño en bits del fichero WAV original y el del fichero OGG, y los valores de PRD se han calculado comparando la señal original y la señal descomprimida, tras haberla convertido a fichero WAV de nuevo.

Calidad	CR	PRD (%)
-0.1	22.5914	20.4117
0.0	20.7209	16.3069
0.1	17.4426	12.8153
0.2	11.7889	10.3200
0.3	7.8286	8.0752
0.4	5.9931	6.6820
0.5	5.1508	5.1868
0.6	4.5986	4.0044
0.7	4.2917	3.2025
0.8	4.0280	2.5237
0.9	3.7728	2.0455
1.0	3.5435	1.5702

Tabla 2. Valores de CR y PRD obtenidos para un registro FCG de un sonido normal de 13.9s de duración, comprimido con el método OGG Vorbis para diferentes calidades de este formato

Este mismo registro también se ha comprimido con el método propuesto en este artículo para $RtEn = 99.9\%$, 99.99% y 99.999% (Tabla 3). Los resultados de los dos métodos, representados en la Figura 5, muestran que la compresión con WT supera claramente a la compresión OGG, obteniendo mayores tasas de compresión con menores distorsiones, lo que corresponde a señales

comprimidas más pequeñas y con menores errores respecto a las señales originales.

Se puede observar que, por ejemplo, para una distorsión de aproximadamente 3.2%, el método propuesto comprime la señal casi 5 veces más que el método OGG (20.9 frente a 4.3). Por otra parte, para una tasa de compresión de aproximadamente 21, nuestro método presenta un error 5 veces inferior al obtenido con OGG (3.1 frente a 16.3).

Finalmente, en la Figura 6 se representan 10 ciclos de la señal que se ha tomado como ejemplo para comparar el método propuesto con el OGG Vorbis, donde se muestran la señal original (FCG), la señal reconstruida (A0), y el error cometido (FCG-A0).

RtEn (%)	CR	PRD (%)
99.900	20.8647	3.1361
99.990	13.7616	1.0643
99.999	10.0894	0.5872

Tabla 3. Valores de CR y PRD obtenidos para un registro FCG de un sonido normal de 13.9s de duración, comprimido con el método propuesto para RtEn = 99.9%, 99.99%, y 99.999%

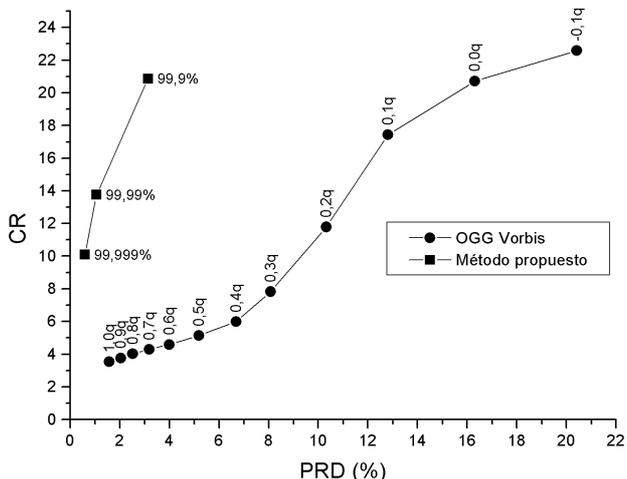


Fig. 5. Representación de los valores de CR frente a los de PRD, tanto para la compresión con OGG Vorbis (indicando la calidad de la compresión) como para el método propuesto (indicando el porcentaje de energía retenida)

4. Conclusiones

En este artículo se ha presentado un método específico para la compresión de señales FCG, basado en técnicas de compresión con pérdidas (aplicación de un umbral a los coeficientes WT y WPT) y de compresión sin pérdidas (codificación RLE y Huffman). Los resultados experimentales han demostrado ser muy superiores a una de las mejores técnicas estándar de compresión de audio, como es el método OGG Vorbis. Esto se debe a que el método propuesto se ha diseñado específicamente para el rango de frecuencias de los FCGs.

Finalmente, queda aún por determinar el error de reconstrucción admisible desde el punto de vista de la percepción del médico, lo que fijará la tasa de compresión obtenida con este método. La aplicación de este método puede ser muy útil en Telemedicina, con el fin de reducir el

ancho de banda necesario para transmitir la señal FCG para análisis o diagnóstico remotos.

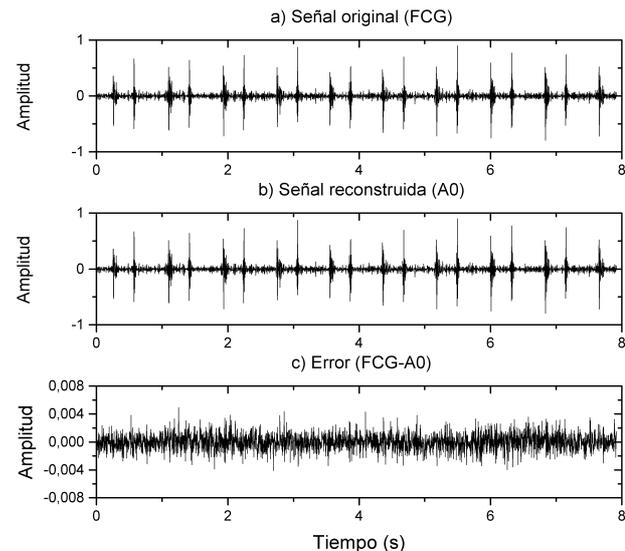


Fig. 6. Representación de a) la señal original (FCG), b) la señal reconstruida (A0), y c) el error (FCG-A0), para una compresión realizada con la WT con los siguientes parámetros: Wmf = Db8, Lb = 4096, RtEn = 99.99%, Nb = 10, y M = 4

Agradecimientos

Este trabajo ha sido financiado parcialmente por la Fundación Séneca de la Región de Murcia y el Ministerio de Ciencia y Tecnología, mediante la concesión de los proyectos PB/63/FS/02 y TIC2003-09400-C04-02, respectivamente.

Referencias

- [1] T. Ölmez, Z. Dokur, "Classification of heart sounds using an artificial neural network", *Pattern Recognition*, 24, 2003, pp. 617-629
- [2] S.M.S. Jalaeddine, C.G. Hutchens, R.D. Strattan, W.A. Coberly, "ECG data compression techniques – A unified approach", *IEEE Trans. on Biomedical Engineering*, 37(4), 1990, pp. 329-343
- [3] B. Bradie, "Wavelet packet-based compression of single lead ECG", *IEEE Trans. on Biomedical Engineering*, 43(5), 1996, pp. 493-501
- [4] M.L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms", *IEEE Trans. on Biomedical Engineering*, 44(5), 1997, pp. 394-402
- [5] B.A. Rajoub, "An efficient coding algorithm for the compression of ECG signals using the wavelet transform", *IEEE Trans. on Biomedical Engineering*, 49(4), 2002, pp. 355-362
- [6] R.R. Coifman, M.V. Wickerhauser, "Entropy-based algorithms for best basis selection", *IEEE Trans. on Information Theory*, 38(2), 1992, pp. 713-718