

Sistema de compresión de la señal ECG basado en la transformada de Karhunen-Loève

Salvador Olmos, José García y Pablo Laguna

Departamento de Ingeniería Electrónica y Comunicaciones. Univ. de Zaragoza.

Resumen

In this work we analyze a Karhunen-Loève transform technique for ECG data compression. This transform has been applied in two different ways: to the entire beat signal and to independent windows (P wave, QRS complex and ST-T complex). The optimum number of coefficients and bits for coding the signal is analyzed for the MIT-BIH Arrhythmia database. The data compression performance of both choices are: for the entire beat a mean compression ratio of 12.1 with a mean MSE of 0.3% and for shorter windows a mean compression ratio of 17.12 with a mean value of MSE of 0.44%.

1 Introducción.

La señal electrocardiográfica (ECG) es una medida no invasiva de la actividad eléctrica del corazón registrada sobre la superficie del cuerpo. Su análisis permite el diagnóstico de muchas cardiopatías, que son una de las principales causas de mortalidad en los países civilizados. Su uso cada vez más extendido como medida de diagnóstico está originando un gran volumen de información para cada paciente. Por ejemplo, para el análisis de arritmias es habitual la utilización de registros Holter de larga duración. Para almacenar uno de estos registros (típicamente de 24 horas de duración, dos canales, con 500 Hz de frecuencia de muestreo y 12 bits de resolución) son necesarios más de 130 MBytes. Por tanto, queda plenamente justificada la necesidad de la compresión de la señal ECG.

La compresión de datos se puede definir como el proceso de detección y eliminación de redundancias en una señal, que en la señal ECG se deben a: un exceso de la frecuencia de muestreo (se muestra en la correlación entre muestras de un mismo latido), a la naturaleza cuasi-periódica de la señal ECG (se muestra en la correlación entre muestras de latidos adyacentes) y a la correlación entre diferentes canales.

Son muchos los métodos de compresión que se han aplicado a la señal ECG. En 1990 Jalaleddine [1] presentó una excelente revisión y clasificó los métodos de compresión en tres grandes familias: *métodos directos*, métodos mediante *transformadas* y métodos de *extracción de parámetros*. Los dos primeros son *reversibles* y de mayor interés ya que permiten la reconstrucción posterior de la señal. Los métodos directos utilizan técnicas de predicción e interpolación, y tienen

un coste computacional bajo, en cambio, su codificación no es capaz de extraer características directamente utilizables por un sistema automático de clasificación. Los métodos basados en transformadas hacen un análisis de la distribución de la energía de la señal en un dominio transformado. Tienen un coste computacional superior a los métodos directos, pero su codificación puede ser utilizada posteriormente por algoritmos automáticos de clasificación.

En este trabajo se propone la aplicación de la transformada de Karhunen-Loève para reducir la correlación entre las muestras de un latido y la codificación diferencial de los coeficientes de la transformada para reducir la redundancia entre muestras de latidos adyacentes. La transformada se aplica de dos formas diferentes: sobre latido completo y sobre segmentos (onda P, complejo QRS y complejo ST-T). A su vez se presenta una comparación con la transformada Discreta Coseno (DCT) aplicada sobre latido completo.

2 Transformada de Karhunen-Loève.

La transformada de Karhunen-Loève (KLT) es la solución del problema de minimizar el error cuadrático medio entre una señal y una combinación lineal finita de funciones ortogonales [2]. La solución de este problema son los autovectores de la matriz de covarianza de la señal, y por tanto se trata de una transformación dependiente de la estadística de la señal.

El cálculo de la transformada de KL comienza por la estimación de la matriz de covarianza de los vectores de señal. La matriz de covarianza \mathbf{C} refleja la distribución de dichos vectores en el espacio de características. Los vectores propios

ortogonales de \mathbf{C} son las funciones base de la transformada \mathbf{T} , y los valores propios λ_i representan la dispersión media de la proyección de los vectores del training set sobre las funciones base correspondientes. Las funciones se ordenarán por orden decreciente de los autovalores para concentrar la energía de la señal en las primeras funciones base.

3 Training set y construcción de las bases de KLT.

Para obtener un training set representativo de señales ECG se ha seleccionado una gran colección de registros de las bases de datos (MIT-BIH Arrhythmia [3] y European ST-T). En primer lugar se detectaron y etiquetaron los complejos QRS utilizando el programa ARISTOTLE [4]. Los latidos se agruparon en tres familias : latidos normales (N), ventriculares (V) y latidos con bloqueo de rama izquierda (L). Para mejorar las prestaciones de la transformada se calcula independientemente una base para cada morfología, ya que se obtienen grupos más homogéneos.

Para el training set de los latidos normales (N) se seleccionaron 114 registros en total (36 de MIT-BIH Arrhythmia y 78 de European ST-T) con un total de más de 220.000 latidos. Para los latidos ventriculares el training set fueron 63 registros (35 de MIT-BIH Arrhythmia, 13 de European ST-T y 5 de MIT-BIH Long-Term) con más de 62.000 latidos. Únicamente 4 registros de las bases de datos contenían la patología (L) con un total de 8.400 latidos.

Para el cálculo de la matriz de covarianza es necesario inventanar y alinear los latidos. La alineación se realiza con respecto a la posición del complejo QRS q_i . Para los latidos normales (N) la ventana de latido completo se define como $(q_i - 250ms, q_{i+1} - 250ms)$. Si el intervalo entre latidos consecutivos $RR_i = q_{i+1} - q_i < 720ms$ la posición de final de latido se calcula como $q_{i+1} + \frac{2}{3}RR_i$. Los latidos con patología (L) se inventanar de forma similar a los normales, pero las cantidades son $380ms$ en lugar de $250ms$, y la fracción $\frac{3}{5}$ en lugar de $\frac{2}{3}$. La ventana de los latidos (V) se define como $(q_i - 0.2RR_i, q_i + 0.8RR_i)$.

Para los latidos (N) y (L) se ha considerado un inventanado adicional para extraer características morfológicas más concretas dentro de cada latido, basándose en la información del origen fisiológico de la señal ECG. Se consideran tres segmentos: onda P ($q_i - 250, q_i - 60$), complejo QRS ($q_i - 60, q_i + 85$) y complejo ST-T ($q_i + 85, q_{i+1} - 250$).

Con el objetivo de eliminar información sin valor clínico antes de calcular la matriz de covarianza se realiza una eliminación de las variaciones de la

línea de base (debidas al efecto de la respiración y movimientos del paciente) mediante interpolación con splines cúbicos. Los latidos se normalizan en amplitud de forma que todos presenten la misma energía y por tanto afecten con el mismo peso en la estimación de la matriz de covarianza de la señal.

Los autovalores de la matriz de covarianza muestran la capacidad de concentración de energía en las primeras funciones base para las señales que componen el training set. El error cuadrático medio en función del número de funciones p utilizados será

$$MSE_p(\%) = \frac{\sum_{i=p}^{P-1} \lambda_i}{\sum_{i=0}^{P-1} \lambda_i} \cdot 100. \quad (1)$$

En la figura 1 se muestra este valor para las bases (N), (L) y (V). Los autovectores con autovalor mayor tienen mayor capacidad de representación de la señal siendo su forma muy similar a la morfología más común de esa familia de latidos. En la figura 2 se muestran los primeros autovectores para la familia de latidos normales (N).

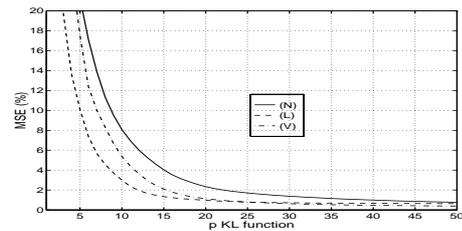


Figura 1: MSE en función del número de funciones base p para (N), (L) y (V).

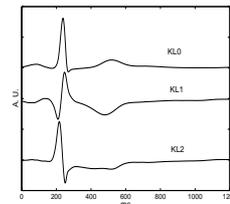


Figura 2: Primeras funciones base (autovectores) de la base de KLT para (N).

4 Compresión de datos mediante transformadas.

El funcionamiento de un sistema de compresión mediante transformadas se refleja en el diagrama de bloques de la figura 3. Cada latido de la señal ECG (vector \mathbf{X} de N componentes enteras codificadas con M bits) se multiplica por la matriz de transformación \mathbf{T} (base ortogonal del espacio vectorial de dimensión N), dando lugar al vector transformado \mathbf{Y} . Las componentes del vector \mathbf{Y} son la magnitud de la proyección del vector \mathbf{X} sobre la base \mathbf{T} . La compresión de datos se puede conseguir seleccionando un reducido número de coeficientes $p < P$, o utilizando un número inferi-

or de bits $m < M$. Se podrá calcular una señal reconstruida X_R aplicando la transformación inversa T^{-1} . La relación de compresión obtenida será $CR = \frac{P \times M}{p \times m}$. Es claro que relación de compresión grande (valores de p, m bajos) conducirá a valores del error de reconstrucción ($X - X_R$) elevados.

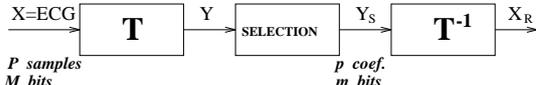


Figura 3: Diagrama de bloques de los sistemas de compresión mediante transformadas.

El sistema de compresión quedará definido por los valores (p, m) que se denota como punto de funcionamiento. La figura 4 muestra las prestaciones de la transformada de KL (MSE) para los latidos normales de los primeros 5 minutos de los registros de las bases de datos estudiadas en todos los puntos de funcionamiento posibles. Se observa como el error cuadrático medio decrece cuando se utiliza mayor número de funciones p y cuando se codifican los coeficientes con un número m mayor de bits. Por analogía a muchas disciplinas, llamamos a estos gráficos curvas características, ya que relacionan los parámetros que definen al sistema (p, m) con sus prestaciones MSE .

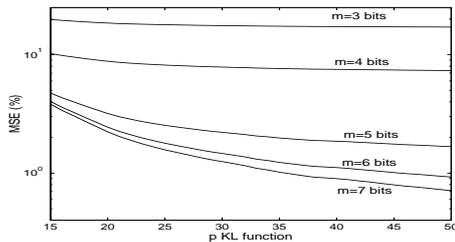


Figura 4: Curvas características para latidos normales (N).

Por ejemplo, si se utilizan $m = 6$ bits y $p = 40$ coeficientes se obtiene un error cuadrático medio MSE de 1% con una relación de compresión de $CR = \frac{P \times M}{p \times m} = 13.2$ para latidos de duración 800 ms ($N=288$ muestras con frecuencia de muestreo 360 Hz) y $M=12$ bits/muestra.

Para obtener el punto de funcionamiento óptimo del sistema (p^*, m^*) es muy útil representar la misma información sobre el plano (p, m) como el mapa de la figura 5. El error de reconstrucción MSE se representa como líneas de contorno continuas. Las líneas discontinuas representan los puntos con la misma relación de compresión, que resultan ser funciones hipérbolas ($p \cdot m = cte$), que implica número de bits/latido (BPB) constante. El valor mínimo de MSE se localiza en la esquina superior derecha, pero también tiene mínima relación de compresión.

Este mapa recoge de forma compendiada el comportamiento de los sistemas de compresión mediante transformadas. En primer lugar, se mues-

tra claramente el compromiso entre relación de compresión y error de reconstrucción: reconstrucciones con bajo error tienen baja relación de compresión y viceversa.

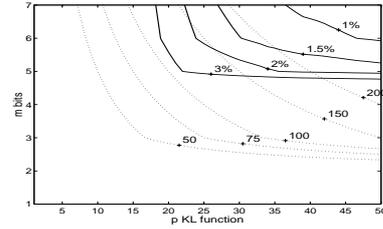


Figura 5: Mapa promedio de error para latidos normales (N).

Para un valor dado de BPB, el punto de funcionamiento óptimo se determina como los valores enteros (p_1, m_1) más próximos a la "rodilla" de la línea de contorno MSE tangente a la hipérbola fijada por BPB. De forma similar, para un valor de MSE dado, el punto de funcionamiento óptimo se determina como los valores enteros (p_2, m_2) más próximos al punto tangente entre dicha línea y una hipérbola.

El método presentado obtiene el compromiso óptimo entre relación de compresión y fidelidad ya que se determina el punto óptimo sobre todo el plano (p, m) . Los mapas de error y sus puntos de funcionamiento óptimos únicamente son válidos para un latido específico. Para representar dicho comportamiento en una población de latidos se considera el mapa promedio de error, como el de la figura 5, que se obtiene promediando los mapas de cada latido.

Los puntos óptimos de funcionamiento para latidos normales con un MSE de 1% se recogen en la tabla 1. También se presenta el punto óptimo para la transformada subóptima DCT, mostrándose la superioridad de la KLT.

MSE	DCT			KL_{beat}			KL_{window}	
	p	m	CR	p	m	CR	Σpm	CR
1%	69	6	7.6	40	6	13.2	185	17.1

Tabla 1: Puntos de funcionamiento óptimos para latidos normales (N).

5 Resultados.

El método propuesto se ha aplicado a registros reales de la base de datos MIT-BIH Arrhythmia Database. El punto de funcionamiento seleccionado es $(p=40, m=6)$ para latidos (N), $(p=23, m=6)$ para (L) y $(p=24, m=6)$ para (V). Los valores de MSE y CR se muestran en las figuras 6 y 7 respectivamente. Se comparan los dos métodos: latido completo y segmentos. El método de latido completo consigue valores de MSE más bajos, pero su relación de compresión es inferior al método de ventanas.

En la figura 8 se muestra un ejemplo de una señal original (primera línea) y su reconstrucción con la base de latido completo (línea intermedia) y la de segmentos (última línea) para el registro 106.

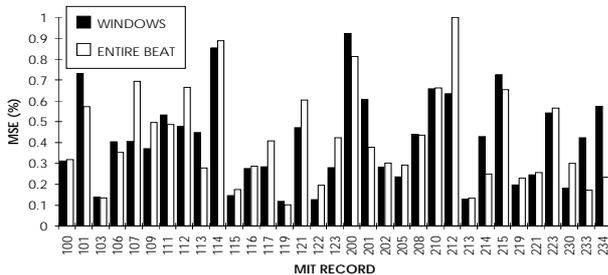


Figura 6: MSE para registros de la base de datos MIT-BIH.

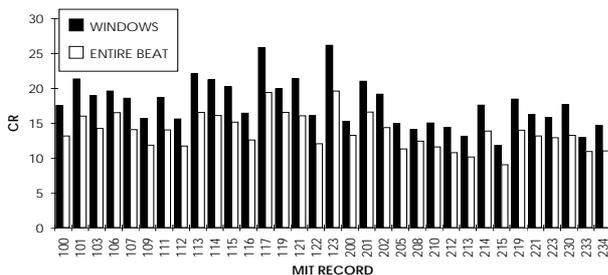


Figura 7: CR para registros de la base de datos MIT-BIH.

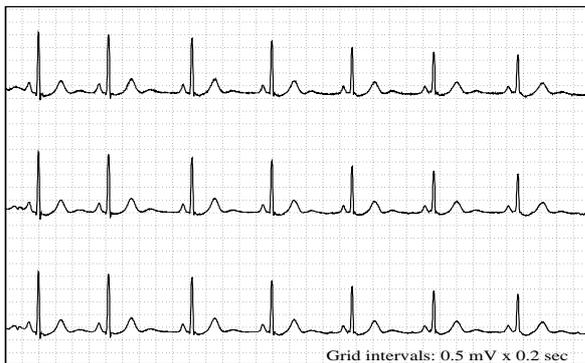


Figura 8: Señal original y reconstrucciones con base de latido completo y segmentos del registro 106.

En la figura 9 se presenta otro ejemplo donde aparecen latidos ventriculares en el registro 221.

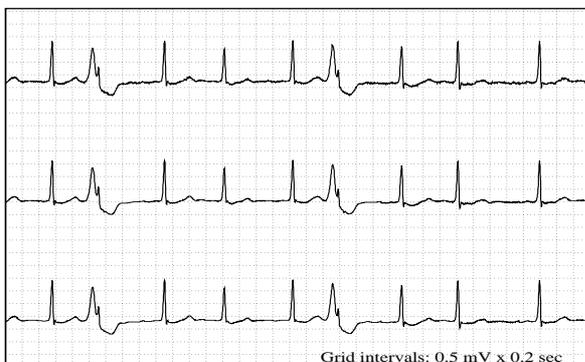


Figura 9: Señal original y reconstrucciones con base de latido completo y segmentos del registro 221.

6 Conclusiones.

En este trabajo se ha aplicado la transformada de Karhunen-Loève para la compresión de señal ECG. A partir de una gran población de latidos se ha estimado una base ortogonal “universal”, en el sentido de que recoge las morfologías más frecuentes. Se han considerado de forma independiente tres familias de latidos: normales (N), ventriculares (V) y bloques de rama izquierda (L).

Para seleccionar el punto de funcionamiento de la transformada (número de funciones p y número de bits m) se confeccionan unos mapas que recogen las prestaciones de todos los posibles puntos de funcionamiento. Las prestaciones de la transformada óptima de Karhunen-Loève se ha comparado con la transformada subóptima Discreta del Coseno (DCT), mostrando la superioridad de la KLT. Además se ha propuesto la aplicación de la KLT a segmentos más específicos del latido (onda P, complejo QRS y complejo ST-T) obteniendo una mejora respecto la aplicación sobre latido completo.

Estos métodos se han evaluado sobre 39 registros de la base de datos MIT-BIH obteniéndose una relación de compresión media de 12.1:1 con un error medio de MSE=0.3% para latido completo y una relación de compresión media de 17.2:1 con un error medio de MSE=0.44% para el método de segmentos.

Agradecimientos

Este trabajo ha sido financiado por el proyecto TIC94-0608-02:01-02. CICYT, y PIT06/93 CONAI.

Referencias

- [1] S. M. S. Jalaeddine, C. G. Hutchens, R. D. Strattan, and W. A. Coberly, “ECG data compression techniques: A unified approach,” *IEEE Trans. Biomed. Eng.*, vol. BME-37, pp. 329–341, Apr. 1990.
- [2] C. W. Therrien, *Discrete random signals and statistical signal processing*. Prentice-Hall, 1992.
- [3] G. B. Moody and R. G. Mark, “The MIT-BIH arrhythmia database on CD-ROM and software for use with it,” in *Computers in Cardiology*, pp. 185–188, IEEE Computer Society Press, 1990.
- [4] G. B. Moody and R. G. Mark, “Development and evaluation of a 2-lead ECG analysis program,” in *Computers in Cardiology*, pp. 39–44, IEEE Computer Society Press, 1982.