

## **DETERMINACIÓN DE INTERVALOS RR A PARTIR DE DATOS CRUDOS DEL ECG**

Dr. Ing. Jesús Rubén Azor Montoya  
Departamento de Investigaciones Científicas y Tecnológicas (DICYT)

CV: Profesor Titular en las asignaturas Estadística Aplicada II y Análisis de Señales en la Facultad de Ingeniería de la Universidad de Mendoza. Ex Investigador en el Instituto de Automática Industrial en Madrid (España). Miembro del Comité del Doctorado en Ingeniería de la Universidad de Mendoza. Expositor en distintos Congresos en EEUU, Puerto Rico, Australia, Sudáfrica, Cuba, etc.

### **Abstract**

La variabilidad del ritmo cardíaco (*Heart Rate Variability*, HRV) describe las variaciones entre latidos consecutivos. El estrés, ciertas enfermedades cardíacas y otros estados patológicos afectan la HRV. Es por ello que en la actualidad se ha puesto mucho énfasis en el análisis de la llamada "variabilidad RR" (donde R es el punto correspondiente al pico del complejo QRS de la onda ECG). Esta variabilidad queda registrada en una serie de tiempos con la cual se procede a aplicar distintas herramientas que permiten extraer información del comportamiento cardíaco. En este trabajo se introduce una técnica sencilla que permite obtener la serie de tiempos de intervalos RR a partir de los datos "crudos" de una salida electrocardiográfica y contrastar los resultados con los registros globales de PhysioBank, una base de datos que contiene más de 36.000 grabaciones de señales fisiológicas digitalizadas organizadas como colecciones de libre disponibilidad.

**Palabras clave:** Series de tiempo, intervalos RR, entropía, filtrado, geometría fractal.

### **1.- Introducción**

El electrocardiograma (ECG) es la manifestación eléctrica de la actividad contráctil del corazón. Se trata de un registro gráfico de la dirección y magnitud de la actividad eléctrica que se genera por la despolarización y repolarización de las aurículas y los ventrículos [Pin11].

Esto proporciona información sobre la frecuencia cardíaca, ritmo y morfología [Tar04]. La importancia de la electrocardiografía es relevante ya que las enfermedades del corazón constituyen una de las principales causas de mortalidad en el mundo.

El desarrollo de la computadora digital y al mismo tiempo la Electrografía ha sido un aliciente en el esfuerzo de los científicos por atender el bienestar humano.

Durante los últimos cincuenta años, mediante la señal del ECG, se han dado pasos importantes en la consecución de sistemas de diagnóstico automatizado que sustituyan a la simple inspección visual.

El ECG varía de persona a persona debido a diferencias en la posición, el tamaño, la anatomía del corazón, la edad, el peso corporal, configuración del tórax y varios otros factores.

El ECG se caracteriza por una secuencia recurrentes de ondas P, QRS, T y U asociadas con cada latido.

El complejo QRS es la forma de onda más destacada y es causada por la despolarización del ventrículo del corazón.

Una típica onda de un latido cardíaco, captado por el ECG de un individuo normal, consiste en una onda P, un complejo QRS y una onda T. La Fig. 1 muestra la forma básica de una señal de latido del corazón sano ECG con P, Q, R, S, J, T y U y características del ECG estándar los intervalos QT, ST y PR.

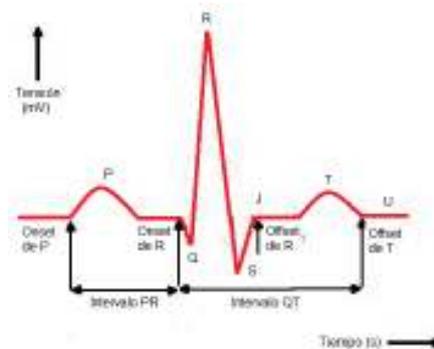


Figura 1 – Señal básica del latido de un corazón sano.

La detección del complejo QRS es uno de los temas fundamentales en el análisis de la señal electrocardiográfica [Tan01]. Consta de tres puntos característicos dentro de un ciclo cardíaco denotados como Q, R y S.

La detección de un complejo QRS no parece ser un problema muy difícil. Sin embargo, en el caso de las señales ruidosas o patológicas o en caso de fuertes variaciones del nivel de amplitud de la misma, la calidad y exactitud de la detección pueden disminuir de manera significativa.

## 2.- Variabilidad de frecuencia cardiaca (HRV)

La *variabilidad de frecuencia cardiaca* (HRV) representa uno de los marcadores más prometedores para medir la actividad del sistema nervioso autónomo - el sistema que es responsable de la mortalidad cardiovascular.

La gran popularidad del estudio HRV está garantizada por la no invasividad y facilidad de obtención, proporcionando una señal invaluable para un análisis profundo del comportamiento cardiaco [Mas09].

El *intervalo RR* es el tiempo transcurrido entre dos ondas R consecutivas en el electrocardiograma, de modo que los datos pueden ser vistos como señales muestreadas con periodo de muestreo no constante, como se aprecia en la figura 2.

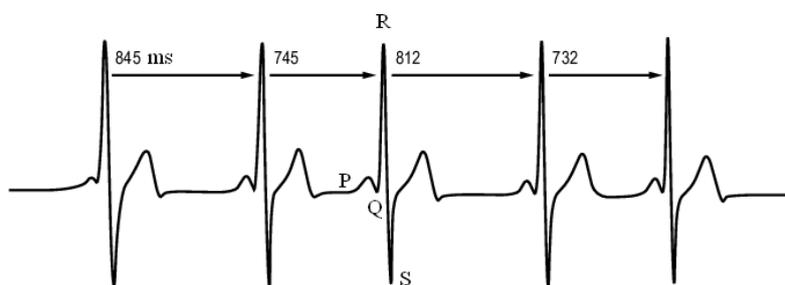


Figura 2 – Cinco latidos cardíacos y cuatro intervalos RR entre ellos.

En adelante, cuando se mencione a la HRV en realidad se hará referencia a la variabilidad de los intervalos RR.

Después de detectarse la aparición del complejo QRS, se puede derivar la serie de tiempo de la HRV. Los intervalos inter-beat o *intervalos RR* se obtienen como diferencia entre los tiempos de ocurrencias sucesivos de la onda R. Es decir, el *n*-ésimo intervalo RR se obtiene como la diferencia entre los tiempos ocurrencia de onda R,  $RR_n = t_n - t_{n-1}$ .

La serie de tiempo construida a partir de todos los intervalos RR disponibles, de muestreo no equidistante, tiene que ser presentada como una función del tiempo, es decir, como valores  $(t_n, RR_n)$ .

El ruido en la serie de tiempo de intervalo RR puede interferir el análisis de estas señales. Las fuentes del mismo se pueden dividir en *técnicos* y *fisiológicos* y debe ser minimizado cuanto sea posible, para lograr análisis confiables.

Otros hechos que pueden alterar significativamente el análisis son tendencias lentas lineales o tendencias más complejas dentro de la serie de tiempo en consideración.

Tales no-estacionariedades lentas son características en las señales de la HRV y deben ser consideradas antes del análisis [Tar02].

PhysioNet ofrece una Web de acceso a una gran colección de registros de señales fisiológicas (<http://physionet.org/physiobank/>) y software open-source (<http://physionet.org/physiotools/>) para la obtención de los datos que habrán de servir como patrón de comparación de la aplicación del algoritmo propuesto en este trabajo.

El "Cajero Automático" de PhysioBank es un centro de auto-servicio para explorar este recurso utilizando un navegador Web. Con esta utilidad se pueden mostrar series de tiempo del intervalo RR e histogramas de las mismas.

En una de sus funciones es posible obtener los intervalos RR entre latidos consecutivos como un archivo ASCII. A modo de referencia se presentan las primeras líneas que se obtienen de consultar el registro "16265" de la base de datos MIT-BIH Normal Sinus Rhythm (nsrdb).

t0	b0	RR (sec)	b1	t1
[08:04:00.000]	[0]	0.406	N	[08:04:00.406]
[08:04:00.406]	N	0.602	N	[08:04:01.008]

Los intervalos RR en sí mismos, aparecen en la tercera columna, flanqueada por los mnemónicos de anotación de los latidos delimitando cada intervalo, y los tiempos de ocurrencia de aquellos latidos.

La versión gráfica de los intervalos en función del tiempo, se puede observar en la Fig. 3.

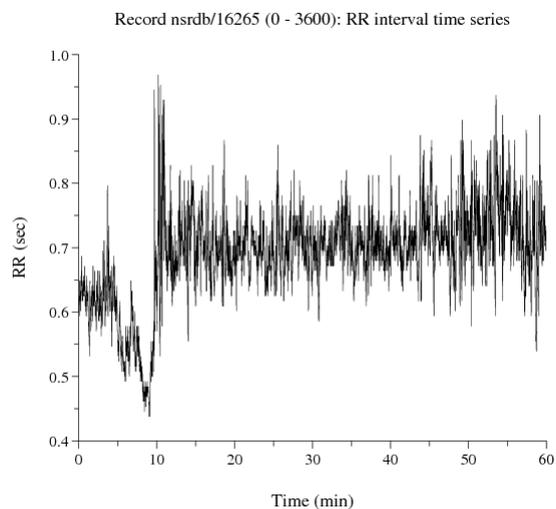


Figura 3 – Gráfico de intervalos RR de una señal de ECG muestreada a 128 Hz en función del tiempo.

### 3.- Información extraída desde los intervalos RR

Una vez lograda la serie de tiempos de intervalos RR, desde los datos del ECG que es el objetivo de este trabajo y se desarrollará más adelante, se habrá de describir algunas de las herramientas utilizadas para la obtención de información de la actividad cardiaca [Tar04, Sha05, Nis01, Mas09].

#### *a – Mapa de Poincaré*

El *mapa de Poincaré*, es un método no lineal basado en el trazado de los intervalos RR en un sistema de dos dimensiones [Bre01]. El método responde a la siguiente definición: dos intervalos RR adyacentes representan un punto del mapa. El primer intervalo RR ( $RR_i$ ) representa la coordenada  $x$ , el segundo intervalo ( $RR_{i+1}$ ) representa la coordenada  $y$ .

La Fig. 4 muestra un mapa de Poincaré de un paciente sano.

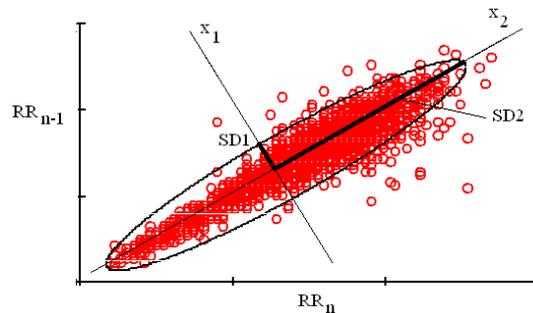


Figura 4 – Mapa de Poincaré de un paciente sano

$SD1$  y  $SD2$  son dos descriptores estándar del mapa de Poincaré.  $SD2$  se define como la desviación estándar de la proyección del mapa de Poincaré sobre la línea identidad ( $y = x$ ), y  $SD1$  es la desviación estándar de la proyección sobre la perpendicular a la línea ( $y = -x$ ).

De la interpretación adecuada de los valores de estos descriptores, se pueden determinar distintas patologías.

#### *b – Entropía aproximada*

La entropía aproximada ( $ApEn$ ) es un estadístico que cuantifica la regularidad en una serie de tiempo. Esta técnica fue propuesta por Pincus [34] como una medida de la complejidad del sistema.

Para su determinación, dada una serie de tiempo de  $N$  datos, se forma una sucesión de  $N-m + 1$  vectores, donde ' $m$ ' es la *dimensión empotrada*. Cada vector consiste de  $m$  puntos consecutivos y además sirve como un vector plantilla para la comparación con todos los otros vectores incluyéndose a sí mismo. Este proceso es referido como self-matching, (auto comparación) en la serie de tiempo, y lleva a la determinación de una probabilidad condicional (condición de que la distancia entre el vector plantilla y los vectores condicionantes estén dentro de la tolerancia ' $r$ ') asociada con este vector.

Finalmente,  $ApEn$  se calcula como el logaritmo natural negativo de la probabilidad mencionada arriba.

Es interesante observar la evolución de  $ApEn$  para un sistema con comportamiento dispar como lo es la ecuación logística [Sha05] que responde a la siguiente función:

$$x[n+1]=r \cdot x[n] \cdot (1-x[n])$$

Este simple modelo es usado frecuentemente para introducir a la *Teoría del Caos* porque visualiza mejor los conceptos caóticos.

La ecuación logística revela una dinámica compleja, dependiendo del valor del parámetro único,  $r$ . La misma ecuación puede generar estados estacionarios, oscilaciones periódicas y un comportamiento muy errático a medida que se produce el aumento de  $r$ .

Para dar una idea acerca de los valores de  $ApEn$  en estas tres situaciones, se lo calcula para valores de  $r$  de 3.55, 3.69, 3.79 arrojando para el estadístico 0, 0.3690, 0.4347 lo que verifica que a mayor caos mayor es el valor de  $ApEn$ .

#### *c – Entropía muestral (SampEn)*

Existe una nueva familia de estadísticos llamada *Entropía muestral (SampEn)* introducida por Richman y Moorman [Ric00], que es un refinamiento de la familia de estadísticos *entropía aproximada (ApEn)* introducida por Pincus. En la entropía muestral, la autocomparación (*self-matching*) que existe en el cálculo de  $ApEn$  se excluye del análisis para reducir el sesgo.

La interpretación de la entropía muestral es similar a la de la entropía aproximada. Para un sistema periódico el valor de  $SampEn$  será muy bajo, mientras que para un comportamiento no lineal determinista y estocástico el valor de  $SampEn$  es alto.

Pincus y Goldberger sentaron la hipótesis de que una amplia clase de enfermedades y perturbaciones representan desacoplamiento y/o reducción de insumos externos del sistema, aislando un componente central desde su universo ambiente.

$SampEn$  proporciona un medio de evaluar esta hipótesis. Un valor alto de  $SampEn$  de la serie de tiempo de intervalos RR corresponde a las condiciones de salud y un bajo valor corresponde a patología.

#### *d – Entropía multi-escala (MSE)*

El análisis *Entropía Multiescala (MSE)* [Cos03, Thu05, Cos07] es un nuevo método de medir la complejidad de las series de tiempo de longitud finita. Esta herramienta computacional se puede aplicar tanto a conjuntos de datos físicos como fisiológicos, y puede utilizarse con una variedad de medidas de entropía.

En lo concerniente al análisis de series de tiempo fisiológicas, la alternativa usada por algunos autores, como Costa et. al., es la *entropía muestral* (*SampEn*).

Las medidas tradicionales de la entropía cuantifican únicamente la regularidad (predictibilidad) de la serie de tiempo en una sola escala. No hay una correspondencia directa, sin embargo, entre regularidad y complejidad.

Ni las señales totalmente predecibles (por ejemplo, periódicas), las cuales tienen entropía mínima, ni las señales completamente impredecibles (por ejemplo, aleatorias no correlacionadas), que tiene la máxima entropía, son verdaderamente complejas, ya que pueden ser descritas muy compactamente. No existe una definición de consenso de la complejidad.

El método *MSE* consiste en someter a la serie de tiempo a un "graneado grueso" construyendo varias sub-series promediando los puntos datos dentro de ventanas no superpuestas de longitud creciente,  $\tau$ . Luego se calcula la *SampEn*, quedando asociada la misma a cada valor  $\tau$ .

La justificación que se hace para abordar este problema [Cos07] comprende tres hipótesis básicas: (i) la complejidad de un sistema biológico refleja su capacidad para adaptarse y funcionar en un entorno en constante cambio; (ii) los sistemas biológicos necesitan operar a través de escalas múltiples tanto espacial como temporal, y por lo tanto su complejidad es también multi-escala y (iii) una amplia clase de estados de enfermedad, así como el envejecimiento, el cual reduce la capacidad de adaptación del individuo, parece que degradan la información transportada por las variables de salida.

Por lo tanto, la pérdida de la complejidad puede ser una característica genérica de la dinámica patológica. En consecuencia, el enfoque para definir una medida de la complejidad se centra en la cuantificación de la información expresada por la dinámica fisiológica en múltiples escalas.

Por lo tanto, el método *MSE* indica que las dinámicas saludables son los más complejas, como se puede observar en la Fig.5 [Cos07].

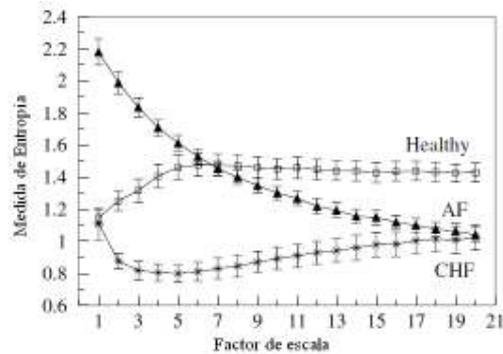


Figura 5 – Análisis MSE de series temporales RR de sujetos sanos en ritmo sinusal normal (*Healthy*), insuficiencia cardíaca congestiva (*CHF*), y con fibrilación auricular (*AF*).

#### 4.- Algoritmo para la elaboración de la serie de intervalos RR

Descrita ya la utilidad de la serie de intervalos RR como una alternativa ampliadora de la información obtenida de los registros electrocardiográficos, se procederá a la determinación de la misma a partir de los "datos crudos" entregados por el ECG.

En primer lugar se determinará el comienzo y fin de cada complejo QRS mediante la aplicación del algoritmo de Pan-Tompkinson [Pan85, Pav03]. Para ello se debe conocer la frecuencia de muestreo de la señal electrocardiográfica ( $F_s$ ) para poder calcular la frecuencia de Nyquist ( $F_{Nq} = F_s/2$ ).

Posteriormente se somete a la señal a un proceso de filtrado, indicado en la Fig. 6.

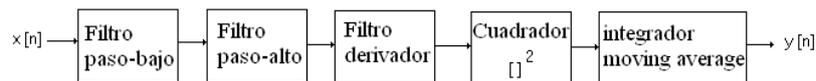


Figura 6 – Diagrama en bloques de la etapa de filtrado.

Los dos primeros bloques representan un filtrado pasa-banda, que reduce la influencia del ruido muscular, la interferencia de 50 Hz, la "baseline wander" (deriva de la línea de base) y la interferencia de la onda T.

La banda de deseable para maximizar la energía del complejo QRS es aproximadamente 5-30 Hz. Se plantea para ello un filtro no-recursivo (FIR) de 130 taps, que tiene la función de transferencia indicada en la Fig. 7.

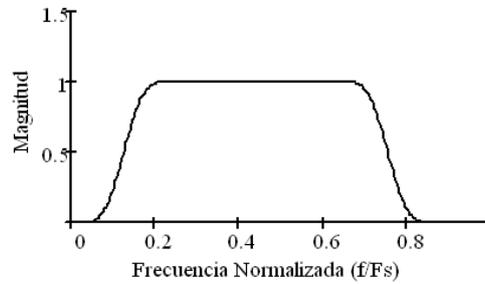


Figura 7 – Respuesta en frecuencia del filtro pasa-banda

Después del filtrado, la señal es diferenciada para proveer al complejo QRS de información acerca de la pendiente. Si se usa un derivador de 5 puntos, la ecuación diferencia tiene la expresión:

$$y[nT] = (2x[nT] + x[nT-T] - x[nT-3T] - 2x[nT-4T]) / 8$$

Posteriormente, la señal es elevada al cuadrado punto por punto. La ecuación de esta operación es:

$$y[nT] = (x[nT])^2$$

Finalmente se procede a la integración *Moving-window* para obtener información de las características de las formas de onda, además de la pendiente de la onda R. Se calcula:

$$y[nT] = (1/N) \cdot (x[nT - (N-1)T] + x[nT - (N-2)T] + \dots + x[nT])$$

El número de muestras  $N$  en la *moving-window* es importante. Generalmente, el ancho de la ventana debería ser aproximadamente el mismo que el complejo QRS más ancho posible. Si la ventana es demasiado ancha, la forma de onda de integración mezclará los complejos QRS y T. Esto puede causar dificultades en posteriores procesos de detección QRS. El ancho de la ventana se determina empíricamente. A modo de ejemplo, para una tasa de muestreo de 200 muestras/s, la ventana es de 30 muestras de ancho (150 ms).

Posteriormente, el algoritmo de Pan-Tompkinson trabaja con una serie de umbrales ajustables que determinan el comienzo y fin del complejo QRS.

Una salida típica de este algoritmo, implementado en Matlab (T.M. The MathWorks, Inc.), se puede ver en la Fig. 8 donde se aprecia la detección del comienzo y fin del complejo QRS.

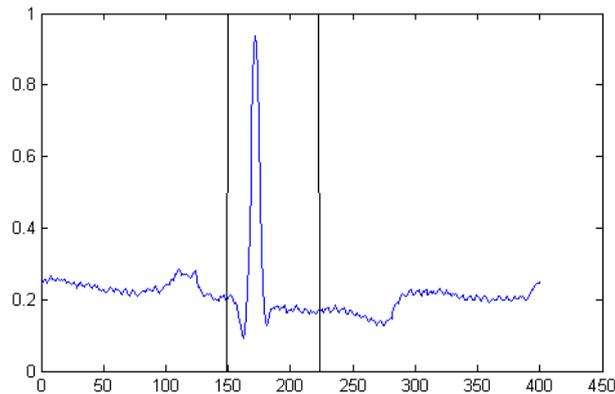


Figura 8 – Determinación del comienzo y fin del complejo QRS de una señal electrocardiográfica muestreada a 360 Hz.

Finalmente, dentro de esas dos acotaciones se determina el valor de abscisa (tiempo) que corresponde al máximo de la señal (punto característico R) quedando establecido un término de la serie de tiempos de intervalos RR.

### 5.- Conclusiones

Por lo anteriormente expuesto, se deduce la importancia de tener el conjunto de datos de los tiempos que separaran a los puntos R del ECG, los que acumulan una preciosa información subyacente complementaria a la que habitualmente maneja un especialista desde la mera señal en el dominio del tiempo.

Todas las nuevas técnicas desarrolladas a través de estas series de tiempos de intervalos RR, como el Mapa de Poincaré, las entropías aproximada, muestral y multi-escala, así como la aplicación de la geometría fractal a través de distintos indicadores [Cip03], dan pie a una ampliación del espectro de herramientas para la detección de patologías.

Con el algoritmo propuesto, se puede suplementar la información suministrada al médico y es posible avanzar fuertemente en la automatización del diagnóstico con simples mejoras en los dispositivos usuales para la adquisición de la señal del ECG [Tis10].

## 6.- Bibliografía

1. [Cip03] CIPRA, B., "*A Healthy Heart Is a Fractal Heart*". SIAM News, Volume 36, Number 7. 2003
2. [Pin91] PINCUS, S. M. , "*Approximate Entropy as a measure of system complexity*". Proc Natl Acad Sci USA, 1991; 88:2297-2301.
3. [Tar04] TARVAINEN, M.P. , "*Estimation Methods for Nonstationary Biosignals*". Doctoral dissertation, Department of Applied Physics University of Kuopio, Finland. 2004.
4. [Mas09] MASEK, O. "*Heart Rate Variability Analysis*". Diploma Thesis, Czech Technical University in Prague. 2009.
5. [Cos03] COSTA, M., GOLDBERGER, A., PENG, C. "*Multiscale Entropy Analysis (MSE)*". Beth Israel Deaconess Medical Center, Boston, USA. Physica A. 2003.
6. [Thu05] THURASINGHAM, R., GOTTWALD, G. "On multiscale entropy analysis for physiological data". School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia. 2005
7. [Tan01] TAN, K. F., CHAN, K. L., CHOI, K. "*Detection of the QRS complex, P wave and T wave in electrocardiogram*". Department of Electronic Engineering, City University of Hong Kong, Hong Kong. 2001.
8. [Bre01] BRENNAN, M., PALANISWAMI, M., KAMEN, P. "*Do Existing Measures of Poincaré Plot Geometry Reflect Nonlinear Features of Heart Rate Variability?*". IEEE Transactions on Biomedical Engineering, Vol. 48, No. 11. 2001.
9. [Pin11] PINZÓN DUQUE, M. C., "*Análisis de señal del impulso cardíaco para el mejoramiento del diagnóstico de patologías del corazón*". Tesis de Maestría en instrumentación Física, Universidad Tecnológica de Pereira (Colombia). 2011.
10. [Ric00] RICHMAN, J.S., MOORMAN, R. "*Physiological time-series analysis using approximate entropy and sample entropy*". Am J Physiol Heart Circ Physiol 278:H2039-H2049. 2000.
11. [Pav03] PAVLATOS, C., DIMOPOULOS, A., MANIS, G., PAPA-KONSTANTINOPOULOS, G. "*Hardware implementation of Pan Tompkins QRS detection algorithm*". National Technical University of Athens, G. Dept. of Electrical and Computer Engineering. 2003.
12. [Tis10] TISMA, R., "*Design of Software for an Electrocardiogram Analyzer*". EE 4BI6 Electrical Engineering Biomedical Capstones. Paper 32. 2010.

13. [Nis01] NISKANEN, J. TARVAINEN, M. RANTA-AHO, P., Karjalainen P. "*Software for advanced HRV analysis*". University of Kuopio Department of Applied Physics Report Series ISSN 0788-4672. Report No. 2/2002. 2002.
14. [Cos07] COSTA, M. PENG, C., GOLDBERGER, A., "*Multiscale Analysis of Heart Rate Dynamics: Entropy and Time Irreversibility Measures*". Springer Science+Business Media, LLC. 2007.
15. [Sha05] SHAH, N. "*Quantification of regularity in RR-interval time series using approximate entropy, sample entropy an multi-scale entropy*". Tesis for Master of Science in Biomedical Engineering. New Jersey's Science & Technology University. 2005.
16. [Ric00] RICHMAN, S., MOORMAN, J. R. , "*Physiological time series analysis using approximate entropy and sample entropy*". The American Journal of Physiology, 2000; 278:H2039-H2049.
17. [Pan85] PAN, J., TOMPKINSON, W., "*A real-time QRS detection algorithm*". IEEE Transaction on Biomedical Engineering, Vol. BME-32, No. 3.1985.