Pre Diagnostician of Heart Arrhythmias Using Wavelet Transform and Fuzzy c Means

Ramón Miranda[1], Isaac Chairez[1], Agustín Cabrera[1]

[1] Unidad Profesional Interdisciplinaria de Biotecnología - Instituto Politécnico Nacional
Av. Acueducto de Guadalupe s/n Col. Ticoman Barrio La Laguna,
CP. 07427, México, D.F.
acabrerall@ipn.mx

Abstract. Heart pathologies are one of the main causes of death in the entire world. Making a correct pre-diagnose can be helpful in order to avoid mistaken treatments that unfortunately are still carried out. For that strong reason, it is presented in this paper, the description of a pre-diagnose process of this heart arrhythmias, using ECG signs and mathematical algorithms as the Continuous Wavelet Transform in order to characterize and identify some features of this pathologies and the Fuzzy c Means that is allowed to identify to what group of this pathologies it belonged an ECG sign of an unknown patient. This Technique could be useful in real medical processes since it can help physicians to carry out a better diagnose.

1 Introduction

1.1 Heart Arrhythmias

In recent years the index of cardiovascular illnesses has been increased, it is caused by inadequate diets, uncontrolled rhythm of life, overweight, or congenital problems among other cases. This information shows the necessity to dedicate more resources to treat and prevent these sufferings by means of improving the diagnosis systems that already exist, which are convenient only for the identification of certain heart pathologies and in spite of the fact that they make this identification with an acceptable error, it is necessary an improvement in the diagnose process.

The current methods, in the pre diagnosis of these pathologies, make an analysis based on the Fourier Transform (FT), however, one of the main inconveniences that this FT has is their non recommended use in non stationary signals. Despite, the normal heart has a non stationary signal it is stiller noted when it presents any kind of arrhythmia. Therefore, in this paper it is suggested to make a pre diagnosis using the Continuous Wavelet Transform (CWT) that has demonstrated better results in certain applications for more than ten years when the signals to be processed are in fact of the non stationary type [5], [9].

© *A. Gelbukh, Á. Kuri (Eds.)*
Advances in Artificial Intelligence and Applications
Research in Computer Science 32, 2007, pp. 419–429

Received 18/06/07
Accepted 31/08/07
Final version 23/10/07

The cardiovascular illnesses are a serious problem of public health in industrialized countries, since they occupy the first place in causes of death, moreover, they are increasing the demand of medical services what generates several socioeconomic repercussions in the health area along the world. For this reason, the use and development of non invasive techniques like electrocardiography (ECG) grants a useful perspective for the diagnosis and treatment of this kind of patients. However, the signal obtained by the ECG is not enough to make a correct diagnose and it is necessary to make a signal processing in order to identify with better results whether a patient has o not any kind of heart pathology [1], [2], [6].

The intervals of time among the different waves of the ECG are important in the diagnosis because they reflect bioelectrical processes that have clinical implications when they are outside of the margin of normal variation. The custom way in which a cardiologist can classify and treat the dysfunctions of the heart rhythm, is by taking into account some ECG characteristic like heart frequency, width of the complex QRS, rhythm state, witness or absence of P waves and their relationship with the complex QRS [3].

1.2 Continuous Wavelet Transform

In the signal analysis there exists a great deal of tools that have been developed for years. Belonging to those tools is the Fourier Transform, the one that has gain considerable attention due to its capacity of represent the frequencies that are contained in a certain signal. The Wavelet analysis allows the use of big intervals of time in those segments where a bigger precision is required in low frequency, and smaller regions where information is required in high frequency [4].

A simple form of understanding the way of operation of this transform is by thinking that the signal in base of time is going to be processed by several low pass and high pass filters which allow the separation of high frequency signal portions from low frequency ones [11], [12].

The objective of the CWT is to decompose certain signals into functions denominated Wavelets components, which form a space base of functions with non finite dimension, and certain properties like orthogonal form, size, softness, duration, etc. The CWT is defined by (1) [7], [8].

$$C(a,b) = \int_{-\infty}^{\infty} f(t)\Psi_{a,b}(t)dt \tag{1}$$

There exist many families into the CWT like Haar, Coifelts and Daubechies. The last one is the used family in this work [9].

1.3 Fuzzy Logic

The human brain carries out a large amount of parallel activities to acquire information about reality, analyze it by means of previous knowledge and to produce an adequate response to the corresponding external excitation. All these processes are developed without mathematical modeling on surrounding systems. This reasoning is main principle of fuzzy logic [10].

In fact, this mathematical theory was developed to reproduce, as far as possible, the human brain capability in order to analyze the real life. The fuzzy logic works with the so-called membership function (MF) as main element to reproduce the uncertainty characteristic on human reasoning. This (MF) concept allows constructing a class of conditions that related the actual particular characteristic on certain process and then to supply a consequent interpretation.

The fuzzy logic methodology introduces the fuzzy groups that relate some particular subsets (MF) defined inside a certain universe of numeric values (representing the real variables involved in the process under study) with some linguistic variables (that are related with the physical meaning of variables nature). This process is known as Fuzzification method, where the real values are evaluated by the MF mathematical representation.

This class of function evaluation is quiet different from the classical Boolean logic, because here the membership degree belong to the closed set (0, 1). Once this process has been completed, a particular methodology is performed to create a conclusion representing an adequate solution to the current situation provided by the actual input variables. Finally, these mathematical expressions are transformed into real physical variables in the Defuzzyfication process.

2 Methodology

In this section it is presented the description of the pre diagnostician process, starting from the data base was used in this work, the signal characterization by the CWT and the clustering by Fuzzy c-Means (FcM).

2.1 Data Base

The database was taken as reference was the MIT - BIH Arrhythmia Dates Base that Contains 48 fragments, corresponding to ambulatory registrations of two channels. The sampling frequency is of 360Hz, with 11 bits of resolution and a range of 10mV. The following table summarizes the characteristics of the used registrations in this work.

Register	*Gender*	*Age*	*Signal Characterization*
101	Female	75	Normal Rythm
104	Female	66	PVC Pacermarker Fusion
105	Female	73	Normal and PVC
107	Male	63	PVC, Paced
109	Masculino	64	Left BBB y PVC
112	Male	54	Normal and APC
200	Masculino	64	High frequency ventricular beat

Table 1. Selected signals of the MIT – BIH arrhythmia data based.

2.2 Mathematical Description of the Wavelet Algorithm

ECG signals obtained by the MIT data base was introduced into a computer. The vector function that represents the bioelectric heart signal is described by (2).

$$\tilde{f} := (x_1, x_2, \ldots, x_n) \tag{2}$$

Every ECG signal is divided into the same number of parts, starting from previous knowledge the size of this intervals was selected as 1 second of acquired signal. The mathematical representation of these intervals is described by (3).

$$\tilde{f} := (\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_k) \tag{3}$$

Every interval is composed by point series (4).

$$\tilde{f}_i := (x_{1i}, x_{2i}, \ldots, x_{\lambda i}) \tag{4}$$

This points are used in order to form a matrix like is shown in (5).

$$F = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{\lambda,1} & 0 \\ \vdots & & & & \vdots \\ x_{1,k} & x_{2,k} & \cdots & x_{\lambda,k} & \end{bmatrix} = \begin{bmatrix} \dot{f}_1 \\ \vdots \\ \dot{f}_k \end{bmatrix} \tag{5}$$

The CWT is applied in every $\dot{f}_i$ and the obtained transform is the expression (6).

$$W_{j,k} = \sum_{\lambda i \in Z} \dot{f}_i[\lambda i]\Psi_{j,k}[\lambda i] \tag{6}$$

Which is a new matrix W_i, shown in (7), that contains the correlated coefficients $c_{f,\lambda k}$ of the corresponding vector $\dot{f}_i$.

$$W_i = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,\lambda 1} \\ \vdots & & \ddots & \vdots \\ c_{f,1} & c_{f,2} & \cdots & c_{f,\lambda k} \end{bmatrix} \tag{7}$$

With the aim of avoid obtain negative coefficients in the process of getting the corresponding energy, it was necessary to obtain the absolute values of the matrix like is described in (8).

$$\hat{W}_i = \left| W_i\{\dot{f}_i\} \right| \tag{8}$$

Finally, it is obtained the proper value of energy of each matrix with the purpose of classify every divided part in frequency and energy. The way was obtained the CWT energy is described in (9).

$$\upsilon_i = \frac{1}{r} \sum_{a=1}^{j} \sum_{b=1}^{r} \hat{W}_i(a,b) \tag{9}$$

Where a is the number of rows and b ius the number of columns.

Once the energy and frequency of every part of the complete ECG signal was obtained, it is possible to visualize different clusters related with every patient condition.

2.3 Mathematical Description of the Fuzzy c-Means Algorithm

The Fuzzy c-Means technique is one of the mostly used clustering tool that has gain considerable attention in real practical applications. This FcM algorithm was developed using an optimization technique in order to solve a minimization problem of a function described by (10).

$$\min_{u_f \times R^{nc}} \left\{ J_m(\mathbf{U},\mathbf{V}) = \sum_{i=1}^{M}\sum_{j=1}^{c}(u_{ij})^m \left\|\mathbf{x}_i - \mathbf{v}_j\right\|^2 \right\} \tag{10}$$

Were $1 < m < \infty$, $V = [\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_c] \in R^{nc}$ and $u_{ij} = u_j(\mathbf{x}_i)$ is the MF which assigns $\mathbf{x}_i$ to the j cluster, and the matrix $M \times c$ established by $\mathbf{U} = [u_{ij}]$ is a partition of FcM given by U_f, which is defined by (11).

$$U_f = \left\{ U \in R^{Mc} u_{ij} \in [0,1],\ \forall i,j;\ \sum_{j=1}^{c} u_{ij} = 1,\ \forall i;\ 0 < \sum_{i=1}^{M} u_{ij} < M,\ \forall j \right\} \tag{11}$$

Then it is defined the group $\Gamma_i = \left\{ j1 \le j \le c : \left\|\mathbf{x}_i - \mathbf{v}_j\right\|^2 = 0 \right\}$ and its complements $\tilde{\Gamma}_i = \{1,2,\ldots,c\} - \Gamma_i$. The cardinality Γ_i of the group Γ_i represents the number of prototypes that coincides with $\mathbf{x}_i$. The necessary conditions to solve $(\mathbf{U},\mathbf{V}) \in U_f \times R^{nc}$ of the minimizing function are shown in (12) if $\Gamma_i \neq \varnothing$, then $u_{ij} = 0$, $\forall j \in \Gamma_i$, y $\Sigma_{j \in \Gamma_i}$, $u_{ij} = 1$, $\forall i$ or (13) if $\Gamma_i = \varnothing$.

$$u_{ij} = \left(\sum_{\varsigma=1}^{c} \left(\frac{\left\|\mathbf{x}_i - \mathbf{v}_j\right\|^2}{\left\|\mathbf{x}_i - \mathbf{v}_\varsigma\right\|^2} \right)^{\frac{1}{m-1}} \right)^{-1}, \quad 1 \le i \le M;\ \ 1 \le j \le c \tag{12}$$

$$\mathbf{v}_j = \frac{\sum_{i=1}^{M}(u_{ij})^m \mathbf{x}_i}{\sum_{i=1}^{M}(u_{ij})^m}, \quad 1 \le j \le c \tag{13}$$

The clusters Fuzyfication produced by the FcM algorithm is controlled by the $m > 1$ parameter.

3 Graphical and Numerical Results

In this section it is presented the obtained graphical and numerical results. Firstly it is presented the ECG signal obtained by the MIT data base, then the graphical results of the CWT application, the relation among frequency and energy and finally the clustering by FcM. Making use of the signals described in the table 1, we analyze every one with this technique. In figure 1, it is presented the signal of a female 75 years old patient with normal ECG signal.

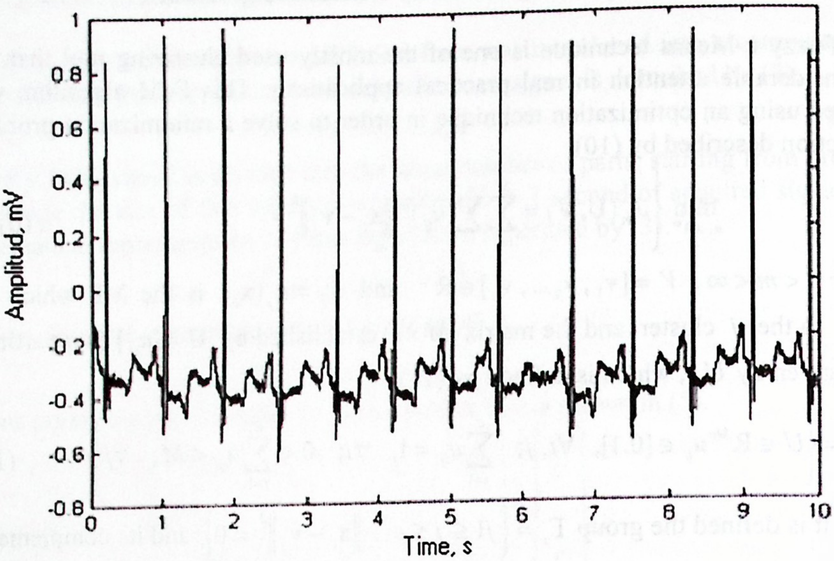

Fig. 1. 100 ECG signal segment.

Once it is applied the CWT the coefficients obtained are 3D plotted and shown in figure 2.

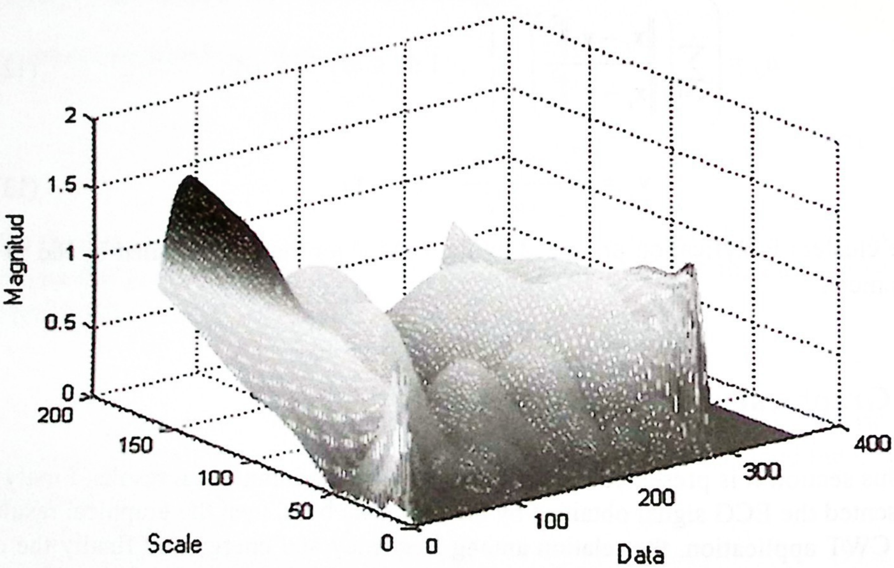

Fig. 2. Coefficients of the 100 ECG signal.

Another ECG signal belonged to a 64 years old male patient with left BBB and PVC condition. This signal is stored in the 109 reference number in the MIT data base. The ECG signal is shown in figure 3.

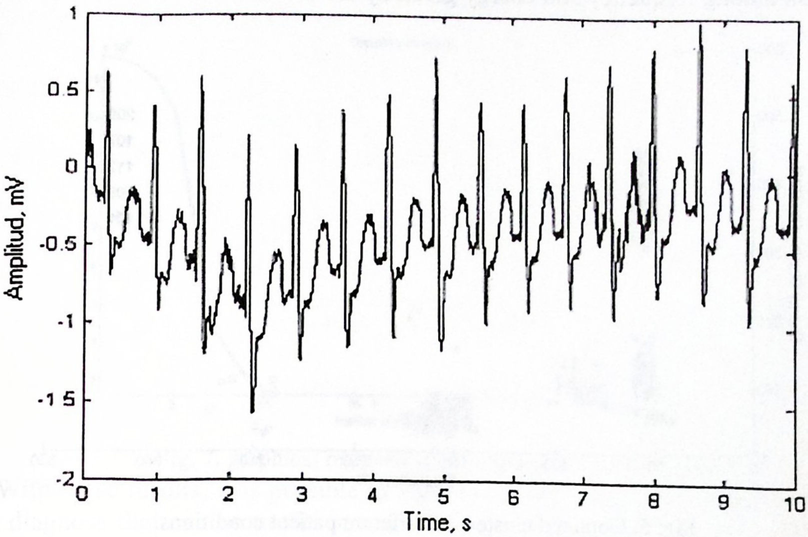

Fig. 3. 109 ECG signal segment.

The CWT application to this signal gives the coefficients that are shown in figure 5.

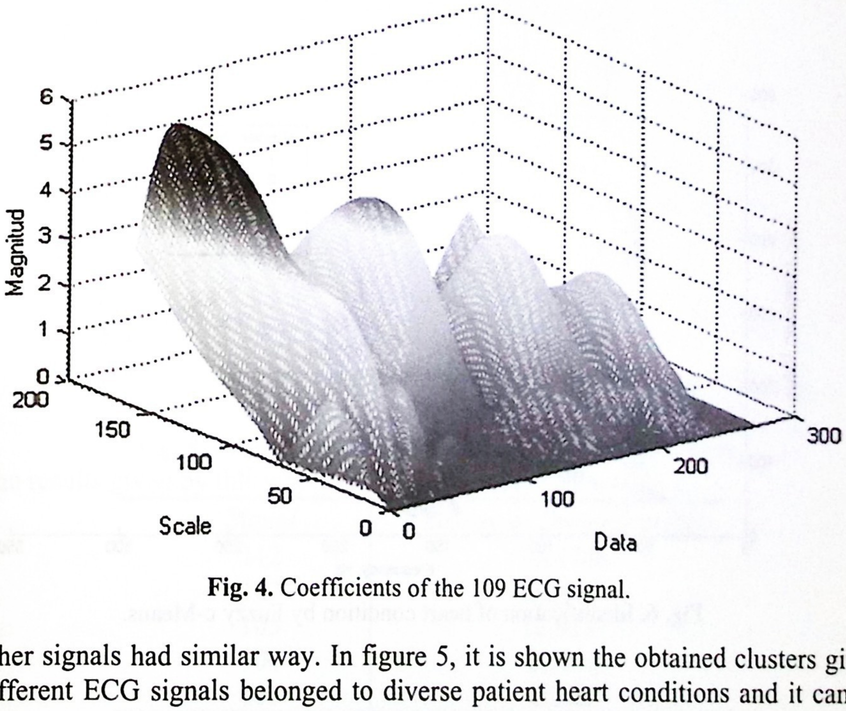

Fig. 4. Coefficients of the 109 ECG signal.

Other signals had similar way. In figure 5, it is shown the obtained clusters given by different ECG signals belonged to diverse patient heart conditions and it can be seen that this algorithm is capable to identify different heart pathologies starting from the relation among frequency and energy given by the CWT.

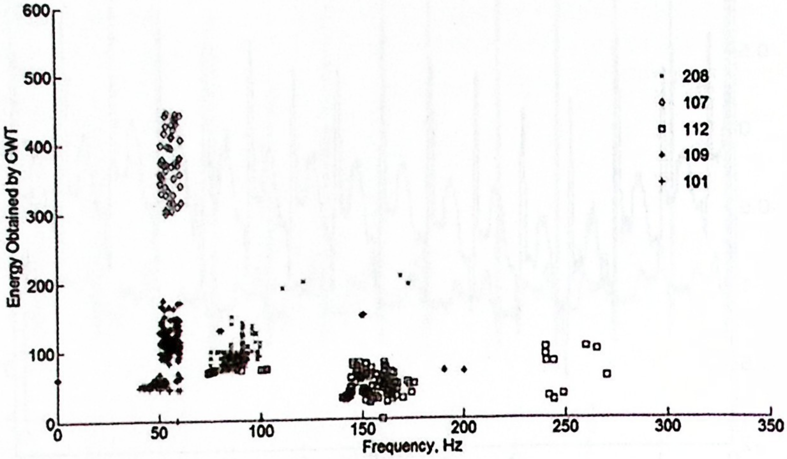

Fig. 5. Obtained clusters of different patient conditions.

Once it is obtained some clusters, the FcM algorithm is applied in order to identify to what group of heart condition is belonged an ECG sign. The obtained results of identification are shown in figure 6 where the continuous lines represent the trajectories of every center given by the FcM technique.

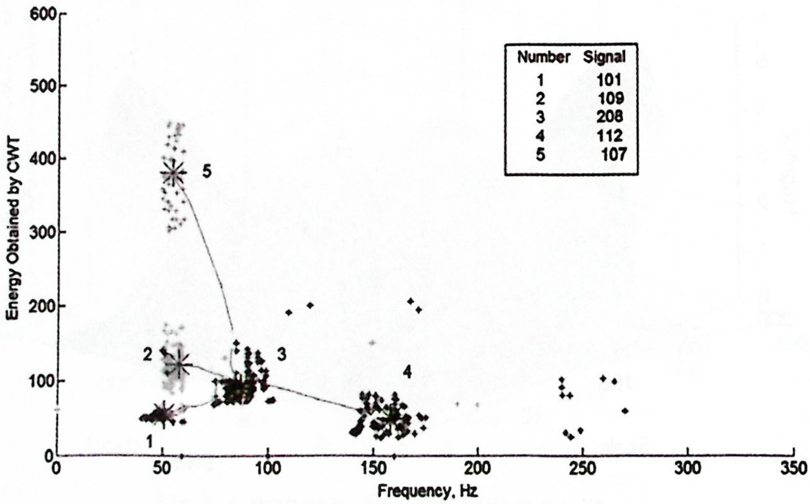

Fig. 6. Identification of heart condition by Fuzzy c-Means.

It is possible to obtain a graphical representation of the minimization problem solution plotting every iteration with its proper value. This evolution is shown in figure 7.

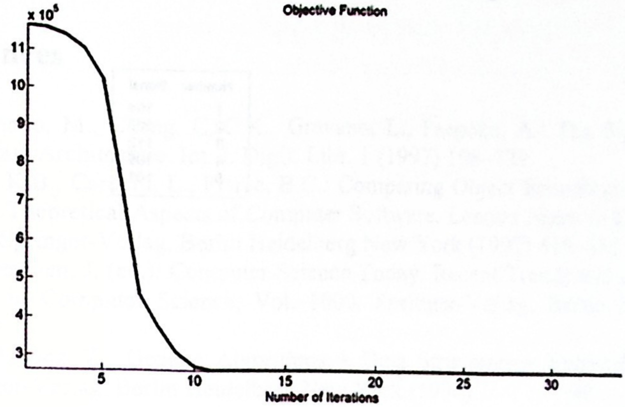

Fig. 7. graphical representation of the minimization solution.

With these results, it is possible to identify an unknown ECG signal and to get a pre diagnose that can be useful in real medical processes. In figure 8 it is presented the relation about frequency and energy of an ECG unknown condition signal and the centers of the characterized known signal given by the FcM technique.

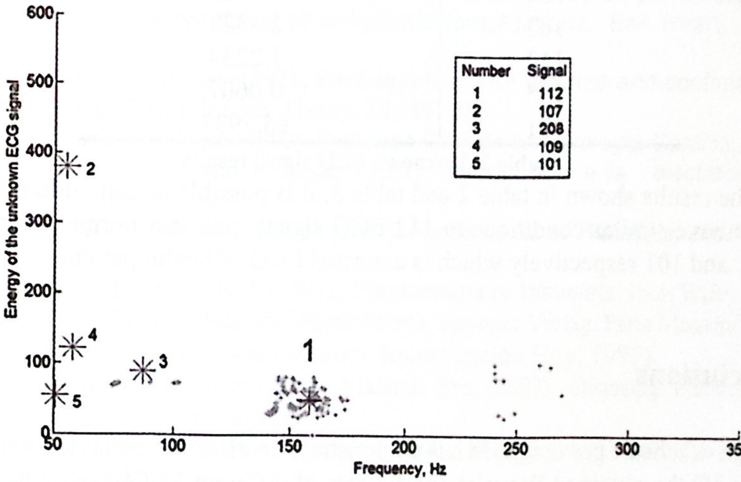

Fig. 8. Unknown ECG signal and the FcM clustering centers.

The results given by this whole technique is presented in table 2.

Signal	*Degree of Ownership [%]*
112	76.4748
107	0.5617
105	12.9137
104	3.7815
101	6.2683

Table 2. Unknown ECG signal results.

Another example of this application is related with the figure 9, where it is presented another different ECG signal of an unknown condition of heart pathologies.

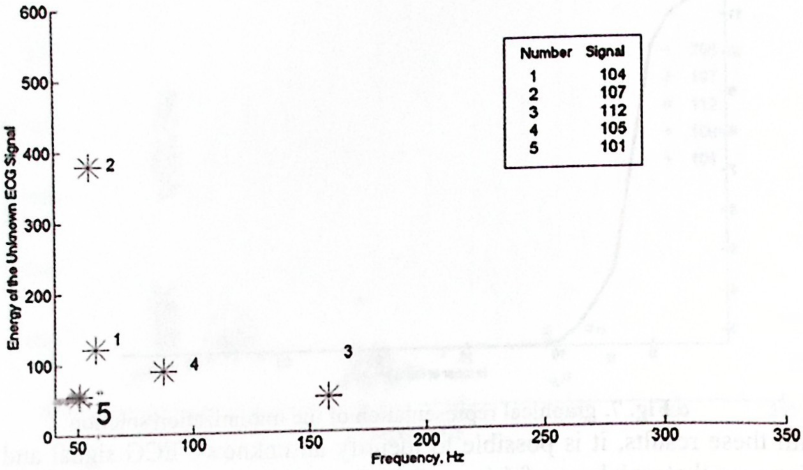

Fig. 8. Unknown ECG signal and the FcM clustering centers.

The obtained results when the FcM algorithm is applied are presented in table 3.

Signal	*Degree of Ownership [%]*
104	0.5903
107	2.3323
112	1.2234
105	0.0607
101	95.7933

Table 2. Unknown ECG signal results.

With the results shown in table 2 and table 3, it is possible to conclude that the patient most have similar conditions to 112 ECG signal such that normal heart rhythm with APC and 101 respectively which is a normal ECG of health patient.

4 Conclusions

The design of a heart pre diagnose has demonstrate an effective behavior while representing in 3D the obtained Wavelet coefficients of different ECG signals belonged to normal and with arrhythmia conditions. It is possible to achieve distinguishing high frequencies and corresponding energies of every signal. With the use of 3D plots obtained with the Continuous Wavelet Transform it is possible to localize crests that corresponds to every wave located in the scale - time plane, furthermore it can be observed that smaller scales correspond to higher frequencies such that frequency diminish when scale grows. Once the CWT is applied, the signal can be represented in 2D frequency - energy plane making clusters depending on the patient condition with the result that the pathologies studied here can be characterized with a representative center using the Fuzzy c-Means technique. In equal manner, this algorithm is suggested to be used in medical applications to analyze different ECG unknown signals in order to help physicians to make a better diagnose.

References

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108–121
3. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415–438
3. van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
4. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)

1. Hsiao-Lung Chan, Shih Chin, Yu-Lin Ko, *Heart Rate Variability Characterization in Daily Physical Activities Using Wavelet Analysis and Multilayer Fuzzy Activity Culstering*, IEEE Transaction on Biomedical Engineerimg, Vol 53, No 1, January 2006
2. Somons, M.B. (1981). Use of signals in the terminal QRS complex to identify patients whit ventricular tachycardia after myocardial infraction. *Circulation* **64**,235-242.
3. Malik, M., O. Odenuyiwa, J. Poloniecki, P. Kulakowski, T. Farrell, A. Stauton and A. J. Camm (1991). *Late potentials after acute myocardial infraction performace of different criteria for the prediction of arrhythmic complications., Eur. Heart. J. 13, 599-607.*
4. Rioul, O. and P. Duhamel (1992)., *Fast algorithms for discrete and continuos wavelet transform*,IEEE Trans. Inform. Theory, **38**,569 - 586.
5. Navarro, F. and F. L. Plaza (1992). *Seciones Clinicas en Patologia Cardica.*, Meditor.
6. Dale, B., U. Dubin and K Linder (2001), *Introducción a la electrocardiografia*, Springer - Verlana Iberica.
7. Jaideva, C., G. Andrew and K Chan (2001). *Fundaments of Wavelets, Jhon Wiley and Sons.*
8. Hartmut, D. and Hartmut H. J. (1997)., *Fundaments of Wavelets*, Jhon Wiley and Sons.
9. Meyer, Y. (1992).*Wavelets and Applications.* Springer Verlag. Paris Masson/ Berlin.
10. V. Ponce, *La Lógica Difusa* (Mexico: Investigación Hoy, 1997).
11. Daniel Novak, David Cuesta Frau, Vladimir Eck (200?); Denosing Electrocardiogram Signal Using Adaptative Wavelets.
12. Nikolay Nikolaev and Atanas Gotchev (200?), ECG Signal Denosing Using Wavelet Domain Wiener Filtering.