ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

Stochastic Method for Non-Homogeneous Cycles Identification in Quasi-Periodic Signals

D. Torres Guzman[1], C. S. Carbajal[2]

PhD Student, Dept of Mechatronics, ITESM-CEM, Atizapan de Zaragoza, Estado de Mexico, Mexico [1]

Professor, Dept of Mechatronics, ITESM-CEM, Atizapan de Zaragoza, Estado de Mexico, Mexico [2]

ABSTRACT: Quasi-periodic signals can be contaminated with random distortions ("artifacts") not manifested periodically and homogenously,without affecting all signal cycles.These distortions cannot be characterized statistically or modelled with a known probability function. In this paper, a stochastic analysis method to detect the presence of such distortions is proposed. The aim of the method is identifying the affected cycles, which exhibit a different morphology compared to the unaffected cycles.The identification of the affected cycles (or non-homogeneous cycles) allows to estimate parameters and extract the useful information needed for a correct characterization of the signal.The method compares nearly periodic signal cycles through the mean square error and the estimated variance of the inherent noise affecting the signal. Expressions are derived to estimate this error and compared with experimental results.

KEYWORDS: Quasi-periodic Signals, Random Distortions,Artifacts, Stochastic Analysis,SignalProcessing, Noise.

I.INTRODUCTION

Measurement, acquisition and processing of quasi-periodic signals, may be affected by many noise sources. On the one hand, there are noise sources generate distortions, which are always present, can be periodic or exhibit a homogeneous behaviour. They affect all signal cycles and are generally modelled as stochastic processes (Gaussian, Poisson, Markov, etc.) with a known probability function using the principles of statistical independence and ergodicity [1]. These distortions include those due to the following noise sources: instrumentation noise(thermal noise, electrostatic noise, electromagnetic noise, processing noise due to stages of analog to digital conversion, environmental noise and interferences) [1-2], improper usage or improper configuration of technology (incorrect colocation of transducers and sensors, erroneous selections of gains,attenuations, filters cutoff frequencies, offset settings etc.) [2] and noise due to own target application (respiratory motions, motions during cycles of systole and diastole of the heart, patient motions, etc.). These types of noise sources have been extensively studied. The papers describing the state of the art includes plenty of methods and systems to their reduction/attenuation.Some of the methods used to attenuate them and to improve de SNR (Signal to Noise Ratio) are:analog and digital classical filtering techniques (high-pass, low-pass and band-pass), advanced filtering techniques, (optimal and adaptive filtering), time-scale and time-frequency transformations, neural networks, averaging techniques, higher order statistics, fuzzy logic, spectral subtraction, linear prediction, Bayesian estimation and many others.For the purposes of this paper, all these noise sources with the above characteristics will be called **inherent noise**.

On the other hand, there are other type of noise sources that produce distortions that are not manifested periodically and homogenously, not affecting all signal cycles.For example: the presence of emboli generates a distortion that will not be distributed at all cycles of the blood flowsignal, the presence of certain cardiac events did not affect all cycles of electrocardiography (ECG) signal, etc.These distortions cannot be characterized statistically or modelled with a known probability function.Such distortions can generate larger deformations to a given segment of the signal, producing a total loss of its morphology, misleading the characterization of the signal under analysis. So, the identification or detection of these types of distortion constitutes one of the most principal challenges in signal processing [3].

The aim of this paper is to propose a method for identifying the cycles of the studied signal which have a different morphology from the rest.In other words, identify those cycles, which, besides being affected by the inherent noise, are affected by distortions that occur on a "casual manner".

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

II.RELATED WORK

One of the most widely used techniques for identifying non-repetitive and/or periodic patterns or distortions has been the wavelet transform [4-5].This technique adapts a wavelet pattern to the characteristics of the signal distortion to be identified.This has been used for identification of epileptic spikes in electroencephalography (EEG) signal [6-9], to identify emboli in the blood flow signal [10-13], to identify arrhythmias in the ECG signal [14-16], for identifying flaws in industrial materials (metals, concrete, etc.)in the ultrasound signal [17-19], and many other scenarios.

The main disadvantage of detection methods based on wavelet is that they depends on the shape and characteristics of pattern to be identified [5], which is not always possible in many real applications.For example, during the verification of a graft, in coronary revascularization procedure, a motion between the ultrasound transducer and the surface (blood vessel 2 to 4 mm in diameter) may cause, in the current acquired cycle, a distortion whose shape is unknown.This distortion may eventually affect one or more signal cycles, and can show different morphology in the affected cycles.This dependence constitutes a limitation for the methods based on wavelets in real-time applications.

The method, in this paper, does not needto identify the distortion characteristics (amplitude, frequency, power, duration, etc.).Indeed, it is independent of the characteristics of the particular application and of the signal carrying useful information [3].

III.PROPOSED METHODOLOGY AND DISCUSSION

A.Quasi-periodic signals and noise sources

The signal $x(t)$is called quasi-periodic with period T if it can be represented, according to a random signal $n(t)$ and to a deterministic and periodic signal $s(t)$, as follows:

$$x(t) = s(t) + n(t)(1)$$

The$k$cycles of the $x(t)$signal, with length T, can be partitioned as follows:

$$x_j(t^{'}) = x(t), \qquad \text{for } t^{'} \in [0, T), t \in [j \cdot T, (j+1) \cdot T), j = 1, 2, 3, \dots, k \qquad (2)$$

$$x_j(t^{'}) = s(t^{'}) + n_j(t^{'}), \qquad \text{for } j \in \{1, 2, \dots k\} \text{and} t^{'} \in \{1, 2, \dots T\}(3)$$

In (3), the signal $s(t^{'})$is invariant cycle to cycle, and it carries the useful information. $n_j(t^{'})$ is associated to the partition, in the cycle j, of the random signal $n(t)$. Also, $n(t)$ will be used to describe the noise sources presented in the measurement, acquisition and processing of the signal $s(t^{'})$. For the purpose of this paper, the random signal $n(t)$ will be divided into two categories. The first category includes the noise sources previously called inherent noise.The set of this noise sources will be represented as n_c. The second category includes the noise sources, which have the following characteristics:theydo not occur in all cycles of the signal and can cause the major deformations in the cycleswhere they occur, including the total loss of their morphology. Furthermore, they are not distributed regularly, they cannot be modeled by a known function probability distribution or statistically characterized, and it is not possible to estimate when or where they will be presented. The set of this noise sources will be represented as n_e.

B. Mean square error

Taken into account the above noise sources, the k cycles of the signal $x(t)$can be represented as follows:

$$x_j = s_j + n_{c,j} + n_{e,j}, \text{for } j = 1, 2, \dots, k \qquad (4)$$

In (4), each cycle has a length of L samples and the deterministic signal s is considered invariant in each cycle (that is, $s_1 = s_2 = \cdots = s_k$). Thus, the mean square error between the k^{th} cycle and the average of the previousacquired cycles is given by:

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

$$.e_k = \frac{1}{L} \cdot \sum_{t=1}^{L} \left[x_k - \frac{1}{k-1} \cdot \sum_{j=1}^{k-1} x_j \right]^2 \quad (5)$$

The noise sources $n_c$are modeled as Gaussian stochastic processes (zero mean) under the principles of ergodicity and statistical independence, therefore, its power in the analysis cycle is equal to its variance; that is, $P_{c,j} = \sigma_{c,j}^2 = \sigma_c^2$, for j = 1, 2, …,k. If the power of noise sources n_e is called $P_{e,j}$, for j = 1, 2, …,k and substituting (4) in (5), it follows that the mean square error is given by:

$$e_k = \frac{k}{k-1} \cdot \sigma_c^2 + P_{e,k} \quad (6)$$

As shownin (6), e_k depends on cycle k under analysis, and if the variance of the noise n_c is known, it is possible to determine if a cycle has been affected by any source n_e (that is, $P_{e,k} \neq 0$), because:

$$e_k > \frac{k}{k-1} \cdot \sigma_c^2 \quad (7)$$

would be satisfied. It should be noted that in each estimation of e_k, using (5), there will be a fluctuation in their own value due to estimation errors.Therefore, when (7) is verified, a region of uncertainty will emerge in (6) where it is not possible to state with certain whether there is any distortion n_e(that is, $P_{e,k} \neq 0$) or an error has occurred due to the estimation of e_k.To explore this uncertainty and the behavior of the estimation of e_k, the random variable E_k is introduced. The variable E_k represents the expected value of $e_k$when $P_{e,k} = 0$ in (6), and is given by:

$$E_k = \frac{k}{k-1} \cdot \sigma_c^2 \quad (8)$$

Since there is some probability that no$n_e$distortions have occurred when $e_k > E_k$, the parameter δ is introduced. Therefore, it will be said that any distortion n_e has occurred if and only if:

$$e_k \geq \delta \cdot E_k \quad (9)$$

The parameter δwill be called **noise factor**and it will control the identification of cycles affected by noisen_e. The product $\delta \cdot E_k$ will be called **detection threshold**and is given by:

$$u_k(\delta) = \delta \cdot E_k = \delta \cdot \frac{k}{k-1} \cdot \sigma_c^2 \quad (10)$$

The choice of parameter δ is critical to the correct functioning of the method. It should be noted that very high values for that parameter are not suitable because the detectionthreshold (expression (10)) will increase. This will cause that fewer affected cycles by noise $n_e$will be detected; so, there would be a significant increase of false negatives (affected cycles that are not detected). Also, very small values (typically close to one) for the parameter δ will not be suitable due to unaffected cycles could be classified as affected; so, there would be a substantial increase in false positives.

C. Noise power estimation

For the identification of the cycles, which are affected by the noise n_e, is necessary to find the detectionthreshold$u_k(\delta)$as shown in (10). For this, is necessary to estimate the variance of the noise sources n_c. Due to the own characteristics of this kind of noise, a good method for achieve that is by using the wavelet decomposition. This method have been extensively used for denoising in many scenarios. In [20-21] the authors demonstrated that the first level decomposition of a signal using wavelet is enough to obtain a good approximation of the noise level in this signal. The
noise level is estimated by the direct relationship between the absolute value of the mean of the detail coefficients and the factor 0.6745.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

D. Selecting the parameterδ

For the purposes of this paper, the selection of parameter δ is not considered crucial. Therefore, the implemented techniques for the selection and adjustment of this parameter will not be exposed here [3]. Furthermore, the values used for δ were selected empirically.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Simulations

To illustrate how the above expressions would be used, a 15-cycles random signal with noise n_e in the cycle 12 is taken as example. The objective is, thus, identify the affected cycle by noise source n_e.Therefore,e_k will be estimate for k = 2, 3,…, 15 using (5), then the noise variance $\sigma_c^2$is estimatedas [20-21] to finally obtain E_k. For a given noise factor value, if (9) is satisfied, then the corresponding cycle could be consider affected by a distortion cause by noise n_e. In Fig. 1 is it shown, with points, the result of $e_k$estimating using (5). The experiment was done over 500 observations of a random signal with 15 cycles of 80 samples each.The experiments were simulated with a Gaussian noise component of variance σ_c^2 in all cycles of the signal. The red dashed line and the blue dashed line,in Fig. 1, represent the E_k valueand the detectionthreshold forδ = 1.5 respectively. In the cycle$k$ = 12,it has been added a noise componentn_e with power $P_e = 1.25 \cdot \sigma_c^2$.

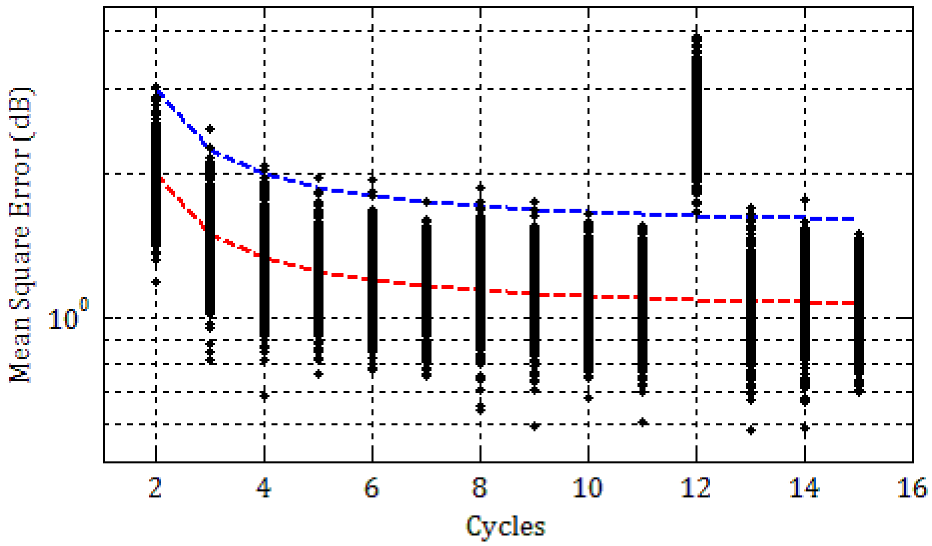

Fig. 1. Representation of the $e_k$estimationfor 500 observations of the 15 cycles of a random signal and noise component n_e with power $P_e = 1.25 \cdot \sigma_c^2$in cycle 12.

It should be noted also in Fig. 1 that there are $e_k$estimates values above the detection threshold (dashed line) in cycles from k = 2 to k = 11 and from k = 13to$k$ = 15. That is, in these cycles, it has been incorrectly detected a noise source n_e(false positive). For cycle k = 12 there are $e_k$estimates values below the detection threshold.In this case we will say that the presence of the noise source n_e(false negative) is not correctly detected. Therefore, it exists an overlap of the affected and unaffected solutions by the noise n_e, which will decrease as the noise power $P_e$increases.
Even with this overlap, it was possible to identify the cycle 12 in Fig. 1 as affected by the noise source n_e, which is the main objective of the proposed method. This detection is useful for eliminating the affected cycle of the signal analysis and of the extraction of important features for the indexes and parametersestimationas will be shown below.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

Real applications

Blood flow signal

The method is applicable over any quasi-periodic signal. Fig.2 shows the envelope of maximum frequencies (normalized with respect to its maximum value) of a Doppler blood flow signal sampled at 11.025kHz.

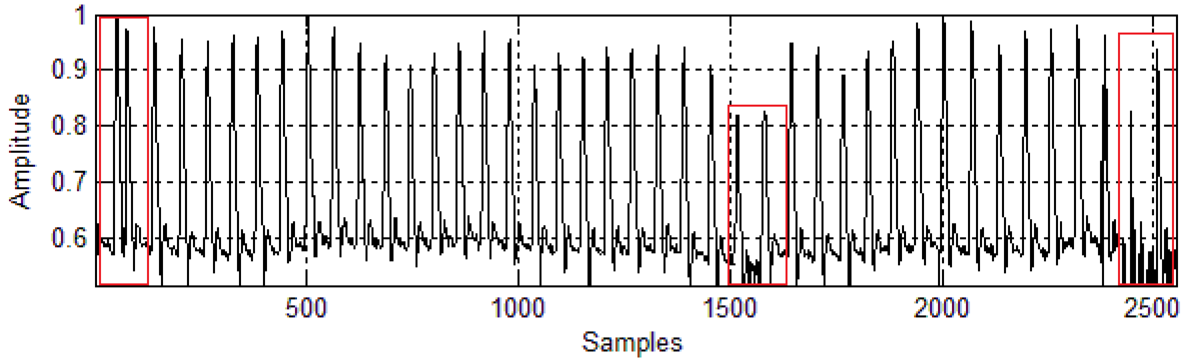

Fig. 2. Envelope of maximum frequencies of a Doppler blood flow signal.

As it can be seen in Fig. 2, there are segments (one or more cycles marked with red rectangles) of the blood flow signal exhibiting a completely different morphology from the rest and they are severely distorted. Affected cycles do not contain clinically useful information. It should be noted that these distortions are not periodic, and they not affect the entire signal (all cycles), therefore we can say that are generated by a noise sourcen_e.The goal is, then, to identify the cycles corresponding to these signal segments. This will allow exclude them of the signal analysisand estimate the clinical indexes needed for a correct diagnosissafely.
For the application of the method, the first step is the delimitation of the signal cycles as indicated by [22]. This is necessary to average the cycles and then estimatethe mean square error between them, as shown in (5). Fig. 3 shows the result of this operation,where the blue vertical dashed lines delimit the signal cycles to a length equal to that of a cardiac cycle.

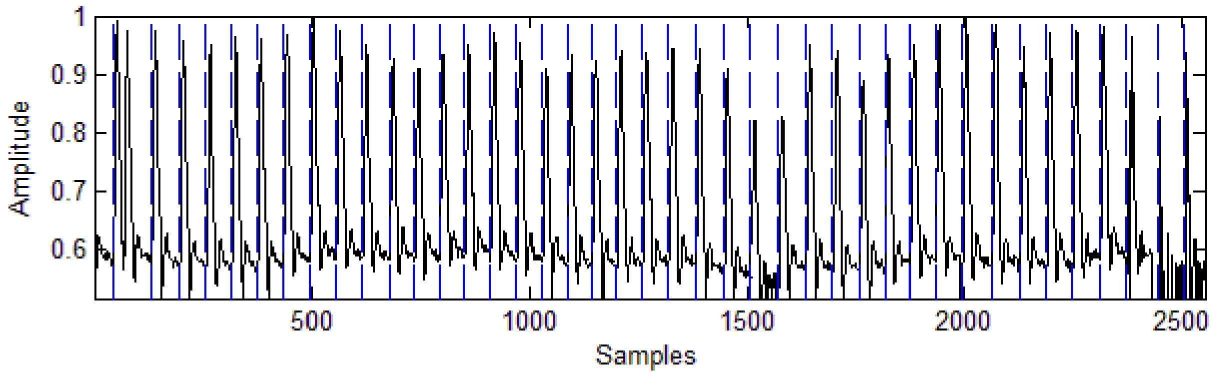

Fig. 3. Envelope of maximum frequencies of a Doppler blood flow signal with cycles delimitation.

Once the cycles has been delimited, the noise varianceis estimated, using wavelet decomposition. Then, a real value (greater than zero) for the parameter δ[3] is proposed and the detection threshold is determined using (10). Finally it is possible identify the no-homogeneous cycles of the analyzed signal estimating the mean square error for each cycle and comparing it with the detection threshold, using (9). The result of all of these operations is shown in Fig. 4.

In Fig.4 the cycles detected as affected by any noise source n_e, are marked with circles and the cycles detected as non-affected are marked with asterisks. The blue dashed line represents the detectionthreshold. It should be noted that,

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

cycles detected as affected match the cycles of the signal shown in Fig. 2 which exhibited a different behavior from the rest.

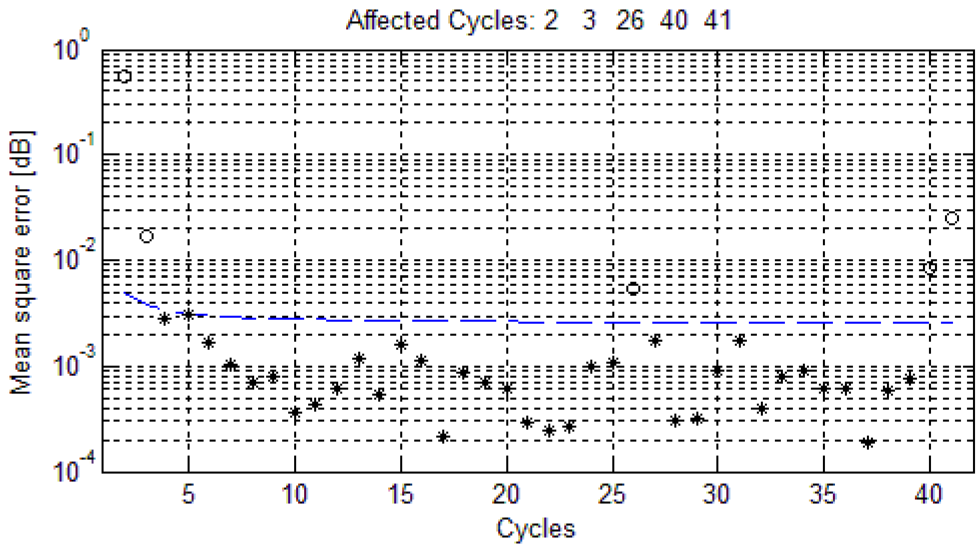

Fig. 4. Classification of the signal cycles as affected (or non-homogeneous) and non-affected by noise n_e.

For a better appreciation of the difference between the cycles classified as affected and unaffected in Fig. 4, the Fig. 5 shows, with the blue line, the final average cycle obtained with the proposed method and two set of cycles; in 5.a the cycles identified as affected (or non-homogeneous cycles) and 5.b the cycles identified as non-affected.

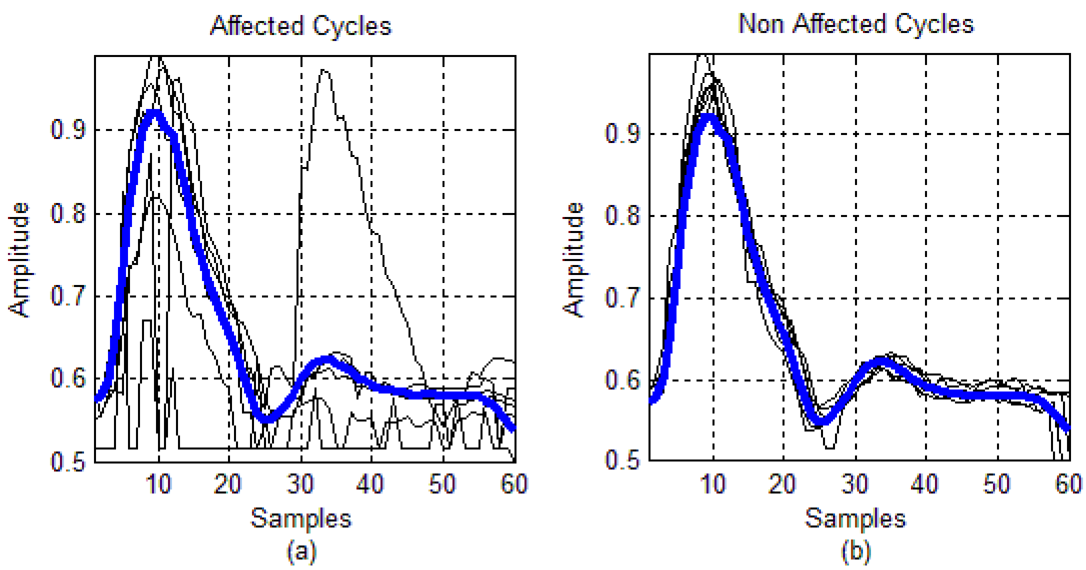

Fig. 5. Final average cycle obtained with the method (blue line). (a) Affected cycles. (b) Non-affected cycles.

As shown in Fig.5, there is a significant difference between the cycles detected as affected and the final average cycle (Fig. 5.a) and between cycles detected as non-affected and the final average cycle (Fig.5.b). Non-homogeneous cycles (Fig. 5.a) show amplitude saturation at the bottom, which distorts the signal information and difficult to extract parameters for the determination of clinical indexesneeded for a correct diagnosis.Depending on the real application, the non-homogeneous cycles (or affected cycles) may be excluded for the estimation of the parameters needed for the clinical indexes computation [3].

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

ECG signal

The Fig.6 shown the 2500 samples of adenoised ECG signal sampled at 250 Hz. In addition, blue vertical dashed lines are used to delimit eachcardiac signal cycle.For a correct delimitation of the cycles, the Gaussian noise elimination techniques was applied firstly.

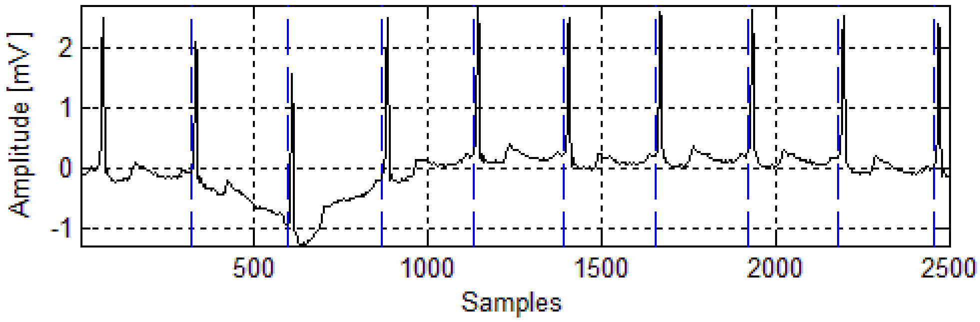

Fig. 6. Denoised ECGsignal with cycles delimitation.

As it can be seen in Fig. 6, the cycles 2, 3 and 4 of the ECG signal exhibit a completely different morphology from the rest. The goal is to identify these cycles, which will reveal the source of noise that originated them.
As in the previous example, once the cyclesare delimited, the noise varianceis estimated, using wavelet decomposition. Then, a real value for the parameter $\delta$is proposed and the detectionthreshold is determined using (10). Then it is possible identify the no-homogeneous cycles of the ECG signal, estimating the mean square error for each cycle and comparing it with the detectionthreshold using (9). The result of all of these operations is shown in Fig. 7.

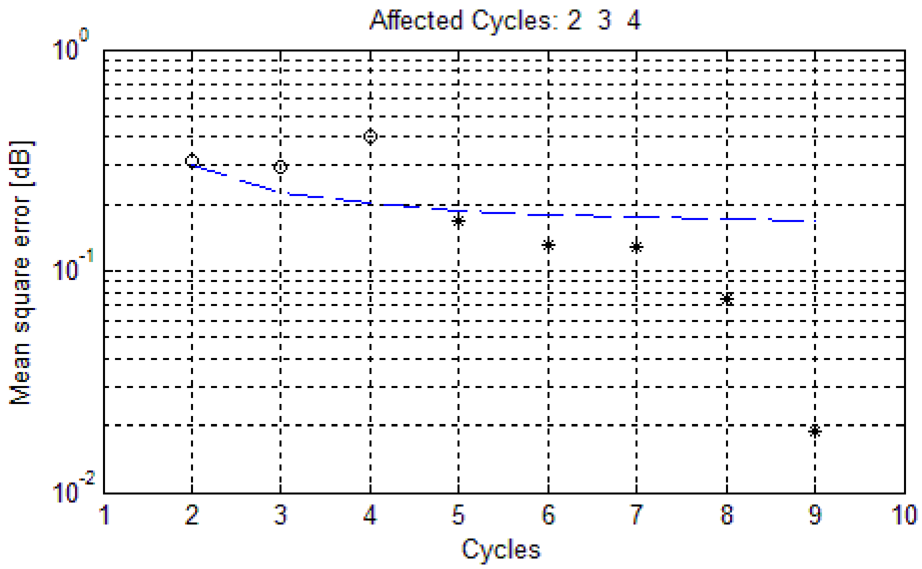

Fig. 7. Classification of the signal cycles as affected (or non-homogeneous) and non-affected by noise n_e.

In Fig.7, the cycles detected as affected by any noise source $n_e$are marked with circles, and cycles detected as unaffected are marked with asterisks. The blue dashed line represents the detectionthreshold. It should be noted that cycles detected as affected match the segment of the signal shown in Fig. 6, which exhibited a behavior different from the rest.
The identification of these non-homogeneous cycles, and its possible exclusion of the signal analysis, is useful for detection of important characteristics (complex QRS detection, arrhythmias, tachycardia, cardiac frequency, heart attack, etc.) for a correct diagnostic of the patient [3].

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2015

V. CONCLUSION

A method for detecting non-homogeneous cycles in quasi-periodic signals was proposed. Such detection is made based on comparisons with a threshold and the estimation of the variance of inherent noise affecting the signal under study. The method is independent of the signal carrying useful information and of the characteristics of distortions to be identified, which is an advantage over the most widely used methods such as those based on wavelets.

The application of the method on two real signals, one of blood flow and other of ECG, was shown. Satisfactory results were obtained and was possible to identify the cycles of these signals which were affected by any noise component $n_e$and do not constitute a source of reliable information for diagnosis.

REFERENCES

1. V. V. Saeed. "Advanced digital signal processing and noise reduction", John Wiley & Sons, Inc., 2nd Ed., New York, USA, 2000. ISBN 0-471-62692-9.
2. S. L. John. "Biosignal and biomedical image processing", Marcel Dekker, Inc., 1st Ed., New York, USA, 2004. ISBN 0-8247-4803-4.
3. D. Torres, C. Carbajal, and J. Isla. "Method to optimize the quality of the blood flow signal by excluding the affected cycles by eventual noise" Patent pending, No. MX/a/2014/014038, 40 p., 2014.
4. S. Pittneret al. "Feature extraction from wavelet coefficients for pattern recognition tasks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 1, pp.83-88, 1999.
5. D. Vinayak, S. Vijay, and M. Mane. "Pattern Recognition using Multilevel Wavelet Transform", International Journal of Computer Applications, Vol. 49, No. 2, 2012.
6. Ganesan M., Sumesh E., and Vidhyalavanya R."Multi-stage, multi-resolution method for automatic characterization of epileptic spikes in EEG", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.3, No. 2, pp. 33-40, 2010.
7. L. Tawade,and H.Warpe."Detection of epilepsy disorder using discrete wavelet transforms using MATLABs", International Journal of Advanced Science and Technology, Vol. 28, pp.17-24, 2011.
8. Haydari Z. et al. "Semi-automatic epilepsy spike detection from EEG signal using genetic algorithm and wavelet transform", IEEE, International Conference on Bioinformatics and Biomedicine Workshops (BIBMW), pp. 635-638, 2011.
9. D. Najumnissa, and T. Rangaswamy. "Detection and classification of epileptic seizures using wavelet feature extraction and adaptive neuro-fuzzy inference system", International Journal of Computational Engineering Research, Vol. 2, No.3, pp. 755-761, 2012.
10. B. Chan et al. "A real-time monitor using wavelet analysis of the Doppler heart sound for the detection of venous air embolism", IEEE International Symposium on Circuits and Systems, pp. 189-192, 1997.
11. N. Aydin, S. Padayachee, H. Markus, "The use of the wavelet transform to describe embolic signals", Elsevier, Ultrasound in Med. & Biol., Vol. 25, No. 6, pp. 953–958, 1999.
12. N. Aydin et al, "Embolic Doppler ultrasound signal detection using discrete wavelet transform", IEEE Transactions on Information Technology in Biomedicine, Vol. 8, No. 2, 2004.
13. N. Aydin. "DWT based adaptive threshold determination in embolic signal detection", IEEE, Second NASA/ESA Conference on Adaptive Hardware and Systems (AHS 2007), 2007.
14. N. Akshay et al. "ECG noise removal and QRS complex detection using UWT", IEEE, International Conference on Electronics and Information Engineering, Vol. 2, pp. 438-442, 2010.
15. V. Ilankumaran, and S. Thamarai Selvi. "Ventricular arrhythmias detection using wavelet decomposition", International Journal of Computer Applications, Vol. 20, No. 1, pp. 11-18, 2011.
16. S. Mukhopadhyay et al. "Wavelet based QRS complex detection of ECG signal", Journal of Engineering Research and Applications, Vol. 2, No. 3, pp. 2361-2365, 2012.
17. Kaya, K et al. "Flaw detection in stainless steel samples using wavelet decomposition", IEEE, Ultrasonics Symposium, Vol. 2, pp. 1271-1274, 1994.
18. Ren-Jean Liou et al. "Flaw detection and sizing of ultrasonic images using wavelet transform and SAFT", IEEE, Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems, pp. 106-110, 2004.
19. E. Oruklu, and J. Saniee. "Ultrasonic flaw detection using discrete wavelet transform for NDE applications", IEEE, Ultrasonics Symposium, Vol.2, pp. 1054-1057, 2004.
20. D. Donoho, J. Johnstone. "Ideal spatial adaptation by wavelet shrinkage". Biometrika, Vol. 81, No. 3, pp. 425–455, 1994.
21. D. Donoho. "De-noising by soft-thresholding," IEEE Trans. on Inf. Theory, Vol. 41, No. 3, pp. 613–627, 1995.
22. D. Evans. "A pulse-foot-seeking algorithm for Doppler ultrasound waveforms". Clinical Physics and Physiological Measurement, Vol. 9, pp. 267-271, 1988.