CLASSIFICATION OF BIOMEDICAL SIGNALS USING THE DYNAMICS OF THE FALSE NEAREST NEIGHBOURS (DFNN) ALGORITHM[1]

Charles Newton Price, Renato J. de Sobral Cintra,
David T. Westwick, Martin Mintchev

Abstract: Accurate and efficient analysis of biomedical signals can be facilitated by proper identification based on their dominant dynamic characteristics (deterministic, chaotic or random). Specific analysis techniques exist to study the dynamics of each of these three categories of signals. However, comprehensive and yet adequately simple screening tools to appropriately classify an unknown incoming biomedical signal are still lacking. This study is aimed at presenting an efficient and simple method to classify model signals into the three categories of deterministic, random or chaotic, using the dynamics of the False Nearest Neighbours (DFNN) algorithm, and then to utilize the developed classification method to assess how some specific biomedical signals position with respect to these categories. Model deterministic, chaotic and random signals were subjected to state space decomposition, followed by specific wavelet and statistical analysis aiming at deriving a comprehensive plot representing the three signal categories in clearly defined clusters. Previously recorded electrogastrographic (EGG) signals subjected to controlled, surgically-invoked uncoupling were submitted to the proposed algorithm, and were classified as chaotic. Although computationally intensive, the developed methodology was found to be extremely useful and convenient to use.

Keywords: Biomedical signals, classification, chaos, multivariate signal analysis, electrogastrography, gastric electrical uncoupling

ACM Classification Keywords: I.5.4 Pattern Recognition: Applications – Signal processing; J.3 Life and Medical Sciences

1. Introduction

Efficient accumulation of accurate knowledge from a wide variety of biomedical phenomena can be obtained from studying and analyzing their dynamics. This dynamics can be assessed by various sensors which monitor, measure, and transform biomedical phenomena into electrical signals that can be analyzed using contemporary electronics and signal processing techniques [1]. Generally, biomedical signals can be to an extent deterministic, random or chaotic [1, 2]. Deterministic signals have the characteristic of predictability, meaning that any future course of the signal could be predicted using some linear analysis tools [1]. Random signals are non-deterministic, in the sense that individual data points of the signal may occur in any order [1], limiting determining the predictability of the future course of the signal to purely stochastic analytical tools. Chaotic signals can be viewed as a connecting mesh between deterministic and random signals, exhibiting behaviour that is slightly predictable, non-periodic, and highly sensitive to initial conditions [2].

Observing that there are three general types of biomedical signals that can be encountered, accurate and efficient study and analysis of these signals can be facilitated by their proper identification as deterministic, random or chaotic, given that specific analysis techniques exist for each type of signals [1]. However, comprehensive and yet adequately simple screening tools to appropriately classify an unknown incoming biomedical signal with respect to these three categories are still lacking.

Recent paper by Gautama et al. [3] presents a method to classify an unknown incoming biomedical signal. The method provides an interpretation of a signal's deterministic and/or stochastic nature in terms of its predictability. Furthermore, it assesses the signal's linear or non-linear nature using surrogate data methods [3]. The result

[1] This study was supported in part by the Natural Sciences and Engineering Research Council of Canada, and by the Gastrointestinal Motility Laboratory (University of Alberta Hospitals) in Edmonton, Alberta, Canada

of this study provides a tool to measure the amount of determinism and randomness in a biomedical signal, useful for detecting a change in health conditions from monitored biomedical signals. However, this method can be seen as an analysis technique that can be applied to a signal once it is classified as deterministic, chaotic or random, rather than as a signal classifier.

The aim of the present work was to develop an efficient and simple method to classify biomedical signals into three categories (deterministic, chaotic or random), using a novel chaos analysis technique which we called the Dynamics of the False Nearest Neighbors (DFNN) algorithm. The proposed method extends the previously developed False Nearest Neighbors (FNN) algorithm [2, 4], to include dynamic FNN characteristics.

2. Methods

Understanding the suggested technique requires an introduction to multivariate signal analysis using state space representation, including time delay and embedding dimension calculations [2, 4, 5].

2.1. State Space Signal Representation

Biomedical signals are usually observed in one-dimensional form, and are represented discretely in the form of a time-domain vector, *s(n)*. It can be inferred that the one-dimensional time-domain vector, *s(n),* is a projection of the signal generator source, represented by an unknown but underlying multidimensional dynamic state vector *x(n)* [2]. The multidimensional dynamic state vector is composed of an unknown number of variables, represented through its dimension *d* [2, 6]. In these notations *n* denotes the current moment in the sampled time-domain.

The transition from a sampled one-dimensional time-domain signal *s(n)* to the corresponding sampled *d*-dimensional state space requires the application of Takens Theorem [6]. Takens Theorem represents a technique to reconstruct an approximation of the unknown dynamic state vector *x(n)* in *d*-dimensional state space by lagging and embedding the observed time series *s(n)*. This reconstructed approximation is the state vector *y(n)* = [*s*(n), *s*(n+T), *s*(n+2T),…, *s*(n+T(*d*–1))], composed of time-delayed samples of *s(n)*, where *T* is the time delay and *d* is the embedding dimension of the system. The accurate calculation of *d* and *T* guarantees through the Embedding Theorem [2], that the sequential order of the reconstructed state vector $y(n) \rightarrow y(n+1)$ is topologically equivalent to the generator state vector $x(n) \rightarrow x(n+1)$, allowing *y(n)* to represent without ambiguity the actual source of the observed multidimensional dynamic vector *x(n)* [2].

Each state space coordinate [s(n), s(n+T), s(n+2T),…,s(n+T(d+1))] constituting a component of *y(n)* defines a point in the state space. As time progresses, the dynamic trajectory of each point in time forms what is called an orbit. An orbit is mathematically defined as the numerical trajectory resulting from the solution of the system [2]. Each orbit constituting *y(n)* is presumed to come from an autonomous set of equations, and therefore, according to the Uniqueness Theorem [2], the trajectory of any orbit is unique and should not overlap with itself. The time delay *T* is an integer multiple of the sampling interval of the signal *s(n)* guaranteeing the extraction of maximal amount of information from the system [2]. The embedding dimension *d* is the minimal state space dimension required to unfold the main orbit of *x(n)* [2]. The main orbit of *x(n)*, known as the attractor, represents the set of points in state space visited by the other orbits of the system long after transients have died out [2].

2.2. Time Delay Calculation

The choice of an accurate time delay *T* guarantees that the time-delayed state space coordinates forming *y(n)* are independent from each other [2]. Choosing too small of a value for *T* clusters the data in state space, while choosing too large of a value for *T* causes the disappearance of the relationships between the points in the attractor [7, 8]. The independence between two coordinates of *y(n)* can be assessed using the mutual information (MI) function [2]. The MI between two *y(n)* coordinates, e.g., *s(n)* and *s(n+T)*, is measured in bits by:

$$MI = log_2 \left\{ \frac{P[s(n), s(n+T)]}{P[s(n)]P[s(n+T)]} \right\}, \quad (1)$$

where $P[s(n), s(n+T)]$ is the joint probability density function (JPDF) of *s(n)* and *s(n+T)*. The average mutual information (AMI) of the JPDFs of all coordinates is calculated by:

$$AMI(T) = \sum_{s(n),s(n+T)} P[s(n),s(n+T)] log_2 \left\{ \frac{P[s(n),s(n+T)]}{P[s(n)]P[s(n+T)]} \right\}. \quad (2)$$

The first minimum of the AMI function provides the optimal time delay T, and assures the independence between the coordinates of the multidimensional vector *y(n)* [2, 7, 8].

2.3. Embedding Dimension Calculation

The signal reconstruction in state space requires a dimension that will guarantee no overlap of the trajectory of the orbit constituting *y(n)*. This optimal dimension is obtained after calculating the percentage of False Nearest Neighbours (FNN) between points in state space. FNNs are calculated using reconstructed state space vectors *y(n)* at different embedding dimensions but a constant time-delay [9]. It is accepted that when the FNN percentage drops to zero, the minimum required dimension to unfold the system into its original state around its attractor is reached, which also guarantees that the orbit is unique [2, 9]. The calculation of the FNNs requires the measurement of a distance R_d, defined as the radius between neighbouring vectors in consecutive dimensions. This procedure is referred to as the FNN algorithm [2, 9]. The square of the Euclidian distance representing R_d as seen in dimension *d* is:

$$R_d(n)^2 = \sum_{m=1}^{d} \left[s(n + T(m-1)) - s^{NN}(n + T(m-1)) \right]^2, \quad (3)$$

where *n* is the current index of the discrete signal (in this case *s(n)*) and s^{NN} is the nearest neighbour (NN) of *s(n)*.

The square of the Euclidian distance in dimension *d+1* becomes:

$$R_{d+1}(n)^2 = \sum_{m=1}^{d+1} \left[s(n + T(m-1)) - s^{NN}(n + T(m-1)) \right]^2 = R_d(n)^2 + \left(s(n + dT) - s^{NN}(n + dT) \right)^2 \quad (4)$$

The change in distance between the points at dimensions *d* and *d+1* is:

$$\sqrt{\frac{R_{d+1}^2(n) - R_d(n)^2}{R_d(n)^2}} = \frac{\left| s(n + dT) - s^{NN}(n + dT) \right|}{R_d(n)}. \quad (5)$$

Determining the existence of a false nearest neighbour depends on how the distance between state space vectors behaves as the calculations progress in consecutive dimensions. If the distance increases significantly with the increment of the embedding dimension, then the vectors are false neighbours, and their closeness results from the reconstruction dynamics of the system, not from its underlying dynamics [2, 9]. If the distance is restricted within a certain threshold level close to the state space points, then the state space points are real neighbours resulting from the dynamics of the system. The embedding dimension that adequately represents the system is the dimension that eliminates most of the false neighbours, leaving a system whose trajectories are positioned in state space due to their underlying dynamics, not to their reconstruction dynamics. Figure 1 shows an example of the results of the FNN algorithm applied to model deterministic, chaotic and random signals.

2.4. Dynamics of FNN (DFNN) Algorithm

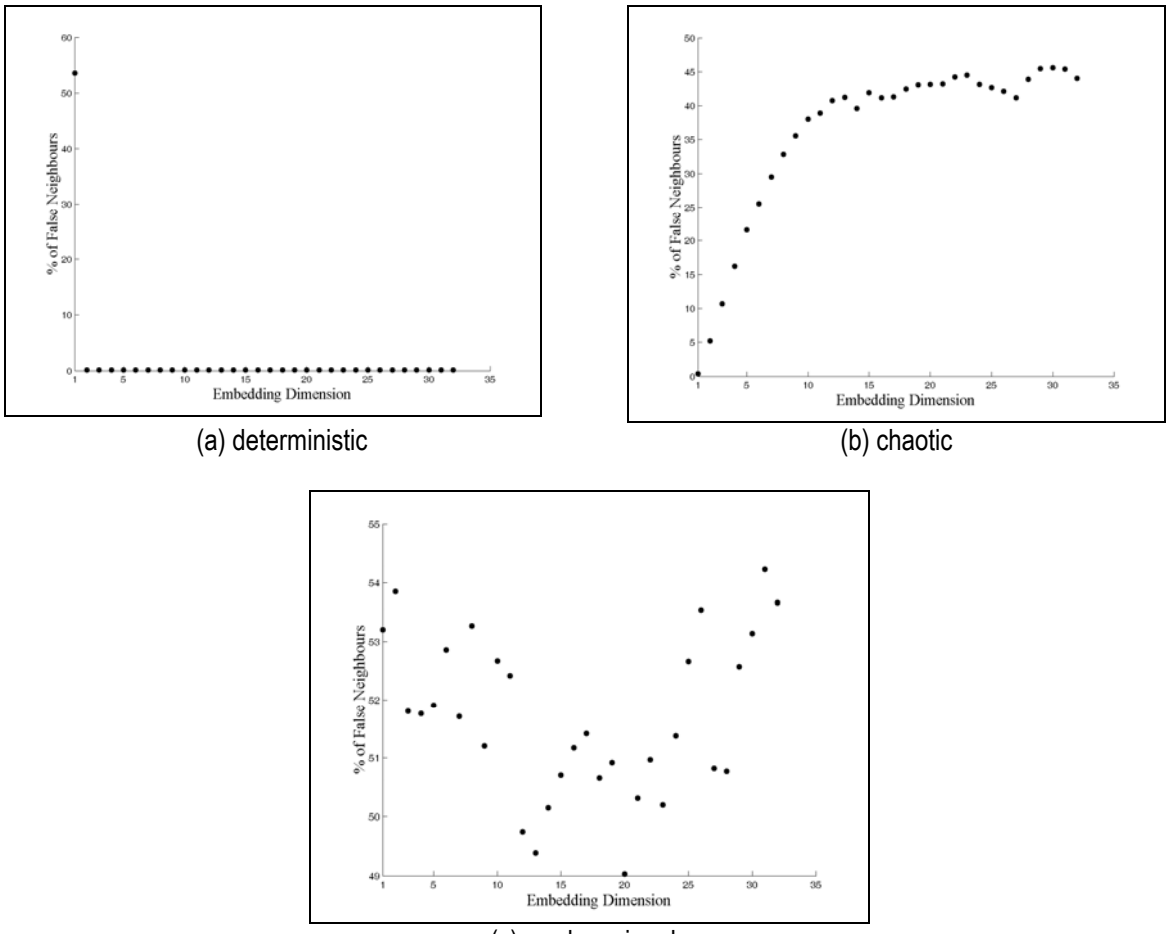
The FNN algorithm is utilized to determine the minimal embedding dimension required to completely reconstruct in state space the source *x(n)* of a one-dimensional time series *s(n)* [2, 9]. Theoretically, the minimal embedded dimension is obtained when the percentage of FNN at a given dimension reaches zero. In practice, not all signals tested through the FNN algorithm reach zero percent FNNs. Two main factors are responsible for this fact. The first is that the larger the embedding dimension of the system, the larger the number of signal samples required for the FNN algorithm [2]. The minimum number of samples required for a given embedding dimension is given by the following equation:

$$m_d = \sqrt{2} \times (e)^d \quad , \tag{6}$$

where *d* is the number of embedding dimensions. The second factor is that an established pattern, or attractor, underlying the dynamics of the system, simply does not exist, as it is the case with white noise [5], for which the FNN algorithm does not reach zero percent FNNs. Therefore, an FNN algorithm failing to converge to zero percent FNN indicates a signal which dimension is too high for the number of samples available.

The important benefit of the proposed DFNN algorithm is that it analyzes the reconstruction dynamics of the signals submitted to the FNN algorithm. Thus, in contrast to the well-established FNN algorithm [9], the DFNN approach does not aim at finding the optimal embedding dimension of the processed signals, but focuses on the processing of the curve representing the FNN dynamics as a function of the embedded dimensions (see Figure 1).

Figure 1 – Sample FNN dynamics of models of (a) deterministic, (b) chaotic, and (c) random signals.



(a) deterministic (b) chaotic

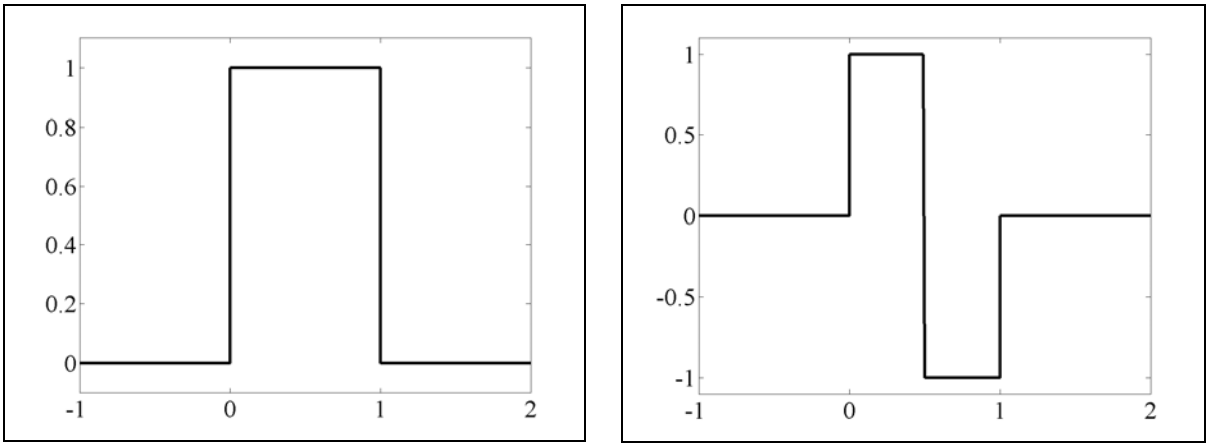
(c) random signals

The independence of the coordinates of the reconstructed state space vector *y(n)* is guaranteed by an accurate choice of time delay *T* [2, 7, 8]. According to the Uniqueness Theorem [2], each reconstruction is unique, therefore the use of an inaccurate embedding dimension when the signal is reconstructed still guarantees a unique representation of the signal, even though it would not represent properly its underlying dynamics until the correct embedding dimension is calculated with the help of the FNN algorithm. What this implies is that the percent of FNNs incrementally calculated at different embedding dimensions in the algorithm is a unique property of the signal under consideration, and therefore, the dynamics of these percent FNNs could also be regarded as a way to represent the signal.

The proposed DFNN algorithm uses a wavelet based pattern recognition technique to classify sampled signals as deterministic, chaotic, or random. The technique is based on analyzing the reconstruction dynamics of the FNNs at consecutively increasing embedding dimensions, ranging from 1 to 32 with the help of wavelet decomposition [10]. The limit of 32 was established due to the fact that 32 is an adequate number of points for a meaningful statistical analysis, because for dyadic wavelet analysis [11] a number that is a power of two is needed, and because it has been shown that the Rossler chaotic system can have an embedding dimension of 25 [12]. It is important to clarify that our aim was not to find the optimal embedding dimension of each signal, but to analyze their reconstruction dynamics as they were submitted to the FNN algorithm.

2.5. Wavelet Decomposition

Wavelet decomposition analysis can be utilized to quantify the shape of the data points extracted from the FNN algorithm [13]. Wavelet analysis coefficients can indicate the resemblance between the shape of a wavelet and a signal. If the resemblance is high, the signal energy is concentrated in few wavelet coefficients. Otherwise, the energy content of the signal is spread throughout these coefficients [14]. Therefore, the aim to quantify the shape of the data points extracted from the FNN algorithm requires (i) finding the wavelet that matches best the wave-shape of the FNN dynamics [14], and (ii) calculating the corresponding wavelet analysis coefficients.



(a) (b)

Figure 2 – (a) Haar scaling function and (b) wavelet.

The proposed DFNN algorithm attempts to classify sampled signals as deterministic, chaotic, or random. An example wave-shape for each signal can be seen in Figure 1. Notice that the shape of the FNN dynamics for each type of signal is different, and that in order to make a comparison one signal type needs to be chosen to be a reference. The deterministic sampled signal pattern was chosen as a reference for determining the wavelet due to its simple shape. The Haar system wavelet was selected, since it matched best the pulse-like wave-shape of deterministic FNN dynamics (compare Figure 1a to Figure 2). Therefore, we hypothesized that through the analysis of the wavelet coefficients, a distinction could be made between deterministic, chaotic and random signal patterns using the FNN dynamics associated with a particular signal.

2.6. Testing the DFNN Algorithm

2.6.1. Categories of Signal Models

In order to test the validity of the proposed DFNN algorithm, signal models with known characteristics were utilized, classified in the existing literature as deterministic, chaotic and random signals [1, 2].

All model signals were sampled frequently enough to comply with the Nyquist Theorem, guaranteeing sufficient digital samples to represent each signal [15]. The length of each signal was 6000 points. The deterministic group of signals included sine, rectangular, triangular, square, sawtooth, quadratic and Dirichlet functions. The models of chaotic signals included the Mackey-Glass Map (MGM), Henon, Ikeda and logistic maps, as well as

the Lorentz, and quadratic systems. The random signal realizations were based on the following statistical distributions: Rayleigh, exponential, beta, χ^2 (chi-squared), gamma, impulsive, normal, and uniform.

2.6.2. Testing Protocol

The three categories of model signals (deterministic, chaotic and random) were submitted to the FNN algorithm, where the percent FNN was calculated and recorded for each embedding dimension up to 32. The resulting 32-point representation of each signal (see samples in Figure 1), were submitted to wavelet analysis decomposition in order to evaluate its corresponding wavelet coefficients. The wavelet coefficients were calculated for each model signal in each corresponding category (deterministic, chaotic and random) and were statistically analyzed by calculating the sample mean and standard deviation. The sample mean and standard deviation were considered distinguishable features to be submitted to the Fuzzy C-means clustering algorithm [16]. The Fuzzy C-means algorithm groups these distinguishable features into clusters. Each cluster provides a centroid, representative of each model signal [16]. This centroid is an important feature corresponding to each cluster that was pivotal for categorizing each signal.

2.6.3. Assessing the Robustness of the DFNN Algorithm

The limited amount of model signals used to test the DFNN algorithm could affect the statistical significance of the results. Therefore, two additional tests were designed to further strengthen the robustness of the DFNN algorithm. The first test involved shuffling the data points of deterministic and chaotic signals to transform them into random signals. These shuffled surrogate signals were then submitted to the DFNN algorithm, and categorized according to their position relative to their nearest cluster using the similarity measure (SM) [17].

The chosen SM is based on the Euclidian metric and is represented by a number between zero and unity, zero representing no similarity, and unity representing maximal similarity. It is calculated by the following equation:

$$SM = \frac{1}{1+l}, \tag{7}$$

where *l* is the Euclidian distance between the centroid of a given cluster (deterministic, chaotic, or random) and the point under consideration.

The second test involved filtering random model signals using a fourth order low-pass digital Butterworth filter. Our hypothesis was that as the random signals undergo low-pass filtering at gradually reduced cut-off frequencies, the level of randomness would be reduced, and the positioning of these new signals in relation to each of the deterministic, chaotic and random centroids, would change after being submitted to the DFNN algorithm. We expected that with the low-pass filtering of these signals with gradually decreasing cut-off frequencies, the level of randomness would drops, and the positioning of the signals would shift away from the random centroid towards the chaotic and deterministic centroids.

2.6.4. Experiment with Electrogastrographic Signals to Detect Gastric Uncoupling

Gastric uncoupling is the loss of electrical synchronization in the stomach [18]. Since gastric motility is electrically controlled, such uncoupling may result in clinical complications such as gastroparesis [18]. In a canine experiment performed by Mintchev et al [18], gastric electrical uncoupling was artificially induced by surgically inhibiting the propagation of electric potentials throughout the length of the stomach using circumferential surgical cuts in the stomach physically separating sections of the organ. Electrogastrography (EGG) is a non-invasive method to record gastric electrical activity [19], and was utilized in this experiment in an attempt to validate its ability to recognize gastric electrical uncoupling. Three kinds of 8-channel EGG signals were recorded from 16 dogs: basal (B), after the first circumferential cut (FC), and after the second cut (SC), each representing three different levels of electrical desynchronization: (i) no uncoupling; (ii) mild induced uncoupling; and (iii) severe induced uncoupling. It has been shown that the amount of electrical uncoupling exhibited in the recorded signals increased with the number of circumferential cuts.

Utilizing the proposed DFNN algorithm, the B, FC, and SC signals from the EGG recordings were tested to assess how each of these signals positions itself with respect to the clusters of deterministic, chaotic, and random signals pre-identified in the experiments with the model signals. We hypothesized that since EGG signals were

found to be chaotic [20], they would position themselves in the chaotic cluster of the plot, and that uncoupling will be detected by noticing that basal EGG signal patterns position closer to the deterministic cluster, while SC signal patterns position themselves closer to the random cluster.

Similarly to the model signals, each of the three types of EGG signals (B, FC, SC) was subjected to the DFNN algorithm. The resulting 128 wave-shape patterns for each state (8 EGG channels per state from each of the 16 dogs) were decomposed using wavelet analysis, and the mean and standard deviation were calculated for the coefficients of each EGG signal type with the aim to show how the B, FC, and SC signals position themselves with respect to the deterministic, chaotic and random regions defined using the model signals. The position of the B, FC, and SC signal patterns in each cluster were quantified using the same SM technique used to test the robustness of the DFNN algorithm [17].

3. Results

3.1. DFNN Algorithm

The percent FNN up to a dimension of 32 were calculated for each model signal using a software package called Visual Recurrence Analysis (VRA) [21]. The calculated dimensions for each model signal resulted in unique wave-shapes (see examples in Figure 1). A total of 49 wave shapes were obtained: 9 from the deterministic model signals, 26 from the chaotic model signals, and 14 from the random model signals. Each of these wave-shapes was submitted to wavelet analysis decomposition using the Haar wavelet, the decomposition resulting in 32 coefficients per model signal. The mean and standard deviation of approximation coefficients per model signal were calculated, with a sample of the results shown in Table 1. A plot of the mean against the standard deviation for each model signal was built (Figure 3). Notice the tendency for the deterministic signals to cluster on the left of the plot, the chaotic signals to cluster near the centre of the plot, and the random signals to cluster to the right of the plot.

To formalize the uniqueness of each group as representative of each model signal, the Fuzzy C-means algorithm was applied to the points of Figure 3. This resulted in a centroid being defined for each model signal group (graphically shown in Figure 3, numerically shown in Table 2), to clearly partition the deterministic, chaotic, and random model signals into specific regions of the plot.

Table 1 – Sample means and standard deviations of the wavelet coefficients obtained from some model signals.

Signal Type	Signal Model	Sample Mean	Sample Standard Deviation
Deterministic	Sine	3.2306	12.4698
	Triangular	2.6224	10.4899
	Square	0.0018	0.0071
	Sawtooth	0.0301	0.0071
Chaotic	Henon	36.7863	11.2378
	Ikeda	55.6975	2.2675
	Logistic	48.7745	19.1534
	Lorentz	40.1871	4.1175
	MGM	25.7374	11.1079
	Quadratic	37.8974	17.3651
Random	Impulse	70.3845	0.5949
	Normal	69.7791	0.6863
	Uniform	70.7354	0.6900

Table 2 - Centroid coordinates for each signal model group.

Signal Type	Centroid coordinate
Deterministic centroid	Mean: 3.3877 SD: 7.2483
Chaotic centroid	Mean: 40.9797 SD: 10.7933
Random centroid	Mean: 69.9643 SD: 1.8843

Table 3 – Means and standard deviations of the similarity measures from all EGG signals.

Signal Type	SM to Deterministic Cluster	SM to Chaotic Cluster	SM to Random Cluster
Basal (B)	Mean: 0.0304 SD: 0.0038	Mean: 0.1329 SD: 0.0449	Mean: 0.0274 SD: 0.0045
First Cut (FC)	Mean: 0.0293 SD: 0.0039	Mean: 0.1549 SD: 0.0586	Mean: 0.0278 SD: 0.0031
Second Cut (SC)	Mean: 0.0298 SD: 0.0039	Mean: 0.1546 SD: 0.0595	Mean: 0.0271 SD: 0.0025

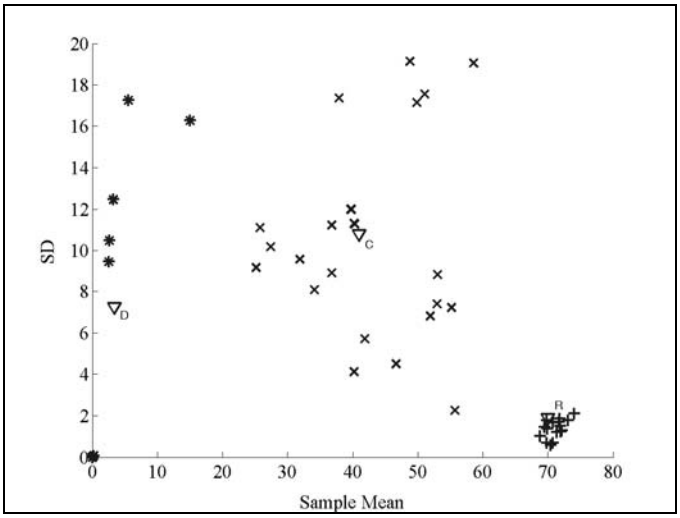


Figure 3 – Sample means and standard deviations of the wavelet coefficients obtained from deterministic (∗), chaotic (×), and random (+) model signals. The centroids (∇) for each cluster were calculated using the Fuzzy C-Means algorithm.

3.2. Robustness of the Algorithm

In the first test performed, the data points of two deterministic and chaotic signals were shuffled and the resulting signals submitted to the DFNN algorithm. All of the shuffled signals positioned themselves in the random region as expected (Figure 4).

The second test for robustness involved filtering random signals using a fourth order low-pass digital Butterworth filter. Three types of random signals were utilized (represented by exponential, uniform and normal probability density functions) and filtered at different normalized cut-off frequencies (0.8, 0.5, 0.3, 0.1, 0.01). With the filtering at decreasing cut-off frequencies, the signals shifted their position further away from the random centroid. This tendency is visualized in Figure 5.

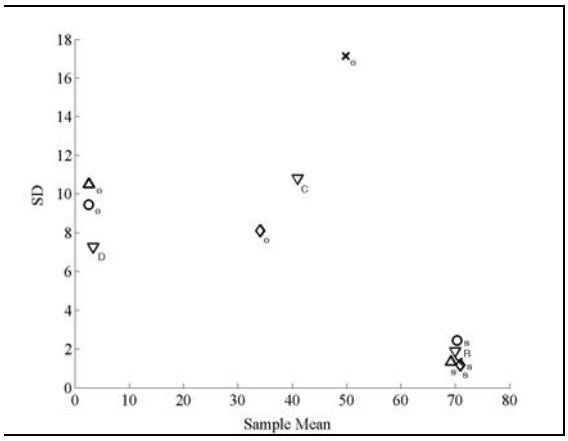


Figure 4 – Positions of the shuffled (s) and the original (o) deterministic (❍ - sine wave, Δ - triangular wave) and chaotic (× - logistic map, ◊ - Lorenz map) signals.

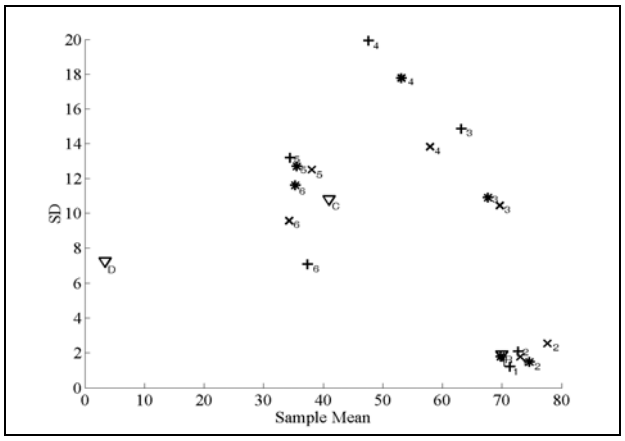
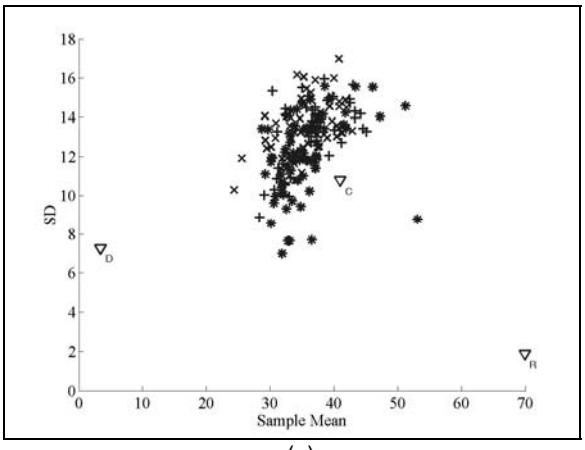


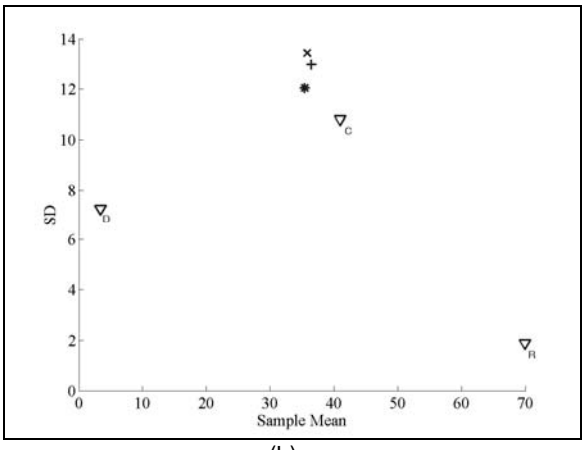
Figure 5 – Three random signals with different probability density functions, exponential (*), uniform (+) and normal (×), filtered with a fourth order low-pass digital Butterworth filter at the following normalized cut-off frequencies: (1) no filtering, (2) 0.8, (3) 0.5, (4) 0.3, (5) 0.2, (6) 0.01.

3.3. Detection of Gastric Electrical Uncoupling

Gastric electrical uncoupling as assessed by the DFNN algorithm can be demonstrated by a single representative point for each EGG signal type, calculated by obtaining the mean of the means and the standard deviations for each of the B, FC and SC signal wavelet coefficients (Figure 6a), resulting in three representative points shown in Figure 6b. Quantitatively, the calculations were performed using Equation 6, where the similarity measure SM was calculated for each of the B, FC, and SC signals with respect to the centroid of each of the deterministic, chaotic and random regions of the plot obtained from the model signals. The overall means and standard deviations for all SM calculations are shown in Table 3. Notice that the similarity of the B, FC and SC signals to the chaotic region was quite strong due to the high SM value, while their similarity to the deterministic and random regions was very weak due to a low SM value. Nevertheless, slight shift was noted towards the random centroid after the first circumferential cut (point FC on Figure 6b), and after the second cut the standard deviation of the obtained wavelet coefficients increased notably (point SC on Figure 6b).



(a)



(b)

Figure 6 (a) Sample mean and standard deviation of the wavelet coefficients from B (*), FC (×), and SC (+) EGG signals.
(b) Positioning of the mean coordinate points representing B (*), FC (×), and SC (+) signals with respect to the centroids (∇) obtained from model signal clusters.

4. Disscussion

In the recent years chaos analysis of biomedical signals evolved into a powerful digital signal processing avenue [2, 4, 5, 22], often overshadowing the well established deterministic and stochastic signal processing tools [1]. However, a comprehensive methodology to determine the adequate digital signal processing tool set (deterministic, chaotic or stochastic) for a specific biomedical signal is still lacking.

In this study we developed an innovative procedure to examine whether given biomedical signals of interest belong to predefined clusters of deterministic, chaotic or random patterns obtained from model signals typical for each of these three categories. The intent was to algorithmically facilitate an informed quantitative decision on which signal processing tools were better suited for the processing of the biomedical signals under consideration. The proposed DFNN algorithm, combined with wavelet decomposition and subsequent statistical analysis were found to be excellent candidates for fulfilling this mission. The research was motivated by the observation that the shapes of the curves produced by the FNN algorithm appeared to be visually related to the signal type. Therefore, a pattern recognition technique based on a dyadic wavelet expansion of the FNN characteristic was developed. The method was tested on a selected set of artificially constructed signals, and then used to assess how some specific biomedical signals [23] position themselves in these categories. It is important to note that the method was tested on a selected set of model signals, and thus further testing with a variety of model signals might be appropriate to fully assess the capabilities and the limitations of the proposed technique.

The methodology resulted in a convenient and very clear clustering of deterministic, chaotic, and random signal patterns extracted from model signals (see Figure 3). Subsequent analysis of electrogastrographic signals in different states (basal, after mild invoked uncoupling, and after severe invoked uncoupling) confirmed previous suggestions that the EGG signals are inheritantly chaotic [20, 24]. Moreover, it was observed that the dominant chaotic nature of these signals, demonstrated by the fact that the DFNN algorithm resulted in their classification well in the middle of the predefined chaotic cluster (see Figure 4), most likely precluded a clear and significant shift from the basal pattern (B) when the signals recorded after the invoked uncouplings (FC and SC) were considered.

5. Conclusion

An innovative technique for classifying biomedical signals in three categories, deterministic, chaotic, and random was developed. The methodology was quantitatively tested using model signals belonging to each of these three categories, and actual electrogastrographic signals subjected to experimentally controlled uncoupling. The technique could be very useful in making an informed decision which digital signal processing toolset would be most appropriate for a specific type of biomedical signals.

Acknowledgement

This study was supported in part by the Natural Sciences and Engineering Research Council of Canada, and by the Gastrointestinal Motility Laboratory (University of Alberta Hospitals) in Edmonton, Alberta, Canada

References

[1] R. E. Challis and R. I. Kitney, "Bio-Medical Signal-Processing (in 4 Parts). 1. Time-Domain Methods," *Medical & Biological Engineering & Computing*, vol. 28, pp. 509-524, 1990.

[2] H. D. I. Abarbanel, *Analysis of observed chaotic data*. New York: Springer, 1996.

[3] T. Gautama, D. P. Mandic, and M. M. Van Hulle, "A novel method for determining the nature of time series," *Biomedical Engineering, IEEE Transactions on*, vol. 51, pp. 728-736, 2004.

[4] B. Chen and N. Wang, "Determining EMG embedding and fractal dimensions and its application," presented at Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE, 2000.

[5] B. Huang and W. Kinsner, "Impact of low-rate sampling on the reconstruction of ECG in phase-space," presented at Electrical and Computer Engineering, 2000 Canadian Conference on, 2000.

[6] F. Takens, "Detecting strange attractors in turbulance," in *Lecture notes in mathematics*, vol. 898, D. A. Rand and L. S. Young, Eds. Berlin: Springer, 1981, pp. 366-381.

[7] W. Liebert and H. G. Schuster, "Proper choice of the time delay for the analysis of chaotic time series," *Physics Letters A*, vol. 142, pp. 107-111, 1989.
[8] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, pp. 1134-1140, 1986.
[9] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical reconstruction," *Physical Review A*, vol. 45, pp. 3403-3411, 1992.
[10] A. M. Reza, "From Fourier Transform to Wavelet Transform, Basic Concepts," Spire Lab, UWM, White Paper October 27 1999.
[11] M. Unser, "Wavelet Theory Demystified," *IEEE Transactions on Signal Processing*, vol. 51, pp. 470 - 483, 2003.
[12] G. Chen, G. Chen, and R. J. P. de Figueiredo, "Feedback control of unknown chaotic dynamical systems based on time-series data," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on [see also Circuits and Systems I: Regular Papers, IEEE Transactions on]*, vol. 46, pp. 640-644, 1999.
[13] A. Ohsumi, H. Ijima, and T. Kuroishi, "Online detection of pulse sequence in random noise using a wavelet," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 47, pp. 2526-2531, 1999.
[14] S. G. Mallat, *A Wavelet Tour of Signal Processing, 2nd ed.*: Academic Press, 1999.
[15] P. P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after Nyquist," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on [see also Circuits and Systems I: Regular Papers, IEEE Transactions on]*, vol. 48, pp. 1094-1109, 2001.
[16] S. Miyamoto, "An overview and new methods in fuzzy clustering," presented at Knowledge-Based Intelligent Electronic Systems, 1998. Proceedings KES '98. 1998 Second International Conference on, 1998.
[17] D. S. Yeung and X. Z. Wang, "Using a neuro-fuzzy technique to improve the clustering based on similarity," presented at Systems, Man, and Cybernetics, 2000 IEEE International Conference on, 2000.
[18] M. P. Mintchev, S. J. Otto, and K. L. Bowes, "Electrogastrography can recognize gastric electrical uncoupling in dogs," *Gastroenterology*, vol. 112, pp. 2006-2011, 1997.
[19] M. A. M. T. Verhagen, L. J. Van Schelven, M. Samson, and A. J. P. M. Smout, "Pitfalls in the analysis of electrogastrographic recordings," *Gastroenterology*, vol. 117, pp. 453-460, 1999.
[20] J. Y. Carre, A. Høst-Madsen, K. L. Bowes, and M. P. Mintchev, "Analysis of the dynamics of the level of chaos associated with gastric electrical uncoupling in dogs," *Medical and Biological Engineering and Computing*, vol. 39, pp. 322-329, 2001.
[21] E. Kononov, "Visual Recurrence Analysis (VRA)," 4.5 ed. http://home.netcom.com/~eugenek, 2003.
[22] I. Yaylali, H. Kocak, and P. Jayakar, "Detection of seizures from small samples using nonlinear dynamic system theory," *Biomedical Engineering, IEEE Transactions on*, vol. 43, pp. 743-751, 1996.
[23] C. Newton Price and M. P. Mintchev, "Quantitative evaluation of the dynamics of external factors influencing canine gastric electrical activity before and after uncoupling," *Journal of Medical Engineering and Technology*, vol. 26, pp. 239-246, 2002.
[24] G. Lindberg, "Is the electrogastrogram a chaotic parameter?," In: Proceedings of the Fourth International Workshop on Electrogastrography, San Francisco, CA, May, 1996.

Authors' Information

Charles Newton Price – M.Sc. graduate student,
Department of Electrical and Computer Engineering; University of Calgary, Calgary, Alberta, Canada T2N 1N4

Renato J. de Sobral Cintra – Ph.D. graduate student,
Department of Electrical and Computer Engineering; University of Calgary, Calgary, Alberta, Canada T2N 1N4
Department of Electronics & Systems, Federal University of Pernambuco; Recife, Pernambuco, Brazil

David T. Westwick – Associate Prof., Dr.,
Department of Electrical and Computer Engineering; University of Calgary, Calgary, Alberta, Canada T2N 1N4

Martin P. Mintchev – Prof., Dr.,
Department of Electrical and Computer Engineering; University of Calgary; 2500 University Drive NW; Calgary, Alberta, Canada T2N 1N4

Department of Surgery, University of Alberta; Edmonton, Alberta, Canada T6G 2B7

phone: (403) 220-5309; fax: (403) 282-6855; e-mail: mintchev@enel.ucalgary.ca