STATISTICAL MODELING OF CHAOTIC ATTRACTORS

Saurabh Prasad

Mississippi State University
Starkville, MS, 39759, USA
email: saurabh.prasad@ieee.org
web: http://www.ece.msstate.edu/

ABSTRACT

A nonlinear dynamical system operating in a chaotic regime is characterized by its sensitive dependence to initial conditions. This leads to state-space trajectories that exhibit locally exponential divergence. Though conventional (linear) signal analysis approaches exploit spectral information contained in the observable, the chaotic structure of an observable generated by a nonlinear dynamical system is best captured by studying the characteristics of its (reconstructed) state-space. This paper reviews some recent advances made by the signal processing community towards capturing a regular, well-behaved statistical structure of an irregular and chaotic deterministic structure. In this paper, I review two such analysis tools. The first one – the Perron-Frobenius operator is used to demonstrate how an irregular evolution in the state-space of a chaotic map settles down to a stable, invariant statistical distribution. Evidence is provided to convince the reader on the possibility of representing the deterministic evolution over the trajectory by a stochastic evolution over a Markov chain. Another useful characterization of some strange attractors is self-similarity on the trajectory. This is characterized by the fractal dimension, and techniques that capture and quantify this property are discussed.

1. INTRODUCTION

A nonlinear dynamical system operating in its chaotic regime will show a state-space evolution that has a strong signature of sensitive dependence to initial conditions. This behavior manifests itself in exponentially diverging trajectories [1] [2] [3] . Dynamical invariants of an attractor capture such signatures (e.g., the largest Lyapunov exponent provides the average rate of exponential divergence of nearby trajectories, the Kolmogorov-Sinai entropy quantifies the rate at which the attractor's trajectory looses / gains information under evolution, etc.).

Conventional (linear) signal analysis tools typically exploit information in the power spectrum of the observable. Since this is based only on second order statistics of the underlying process, such a representation may be inadequate for observables generated by a process with inherent nonlinearities. It is hence desired to use analysis techniques that capture the nonlinear evolution of the underlying system's attractor which generated the observable.

Recent studies suggest that the complicated, deterministic evolution of chaotic attractors settles down to a well-defined statistical structure of the state-space. It has been proven (e.g., using the Perron-Frobenius operator) that under certain conditions, well-known chaotic maps settle down to invariant probability density functions, that can be obtained in a closed form. In this paper, I review the use of some popular statistical methods for modeling trajectories of chaotic attractors. For observables known to have been generated by a deterministic structure, a state-space reconstruction, followed by statistical analysis in this space should provide useful statistical information for modeling purposes. However, it is well known that systems with stochastic differential equations of a fractal order are also capable of generating signals that exhibit strange behavior (e.g., a fractal or self-similar behavior). Clearly, such signals can not be modeled using purely deterministic tools. Recently, the use of fractal geometry for modeling such signals has been explored in the signal processing community. Hence, at a top level, any signal analysis design to capture the nonlinear dynamical structure of the observable must be a two pronged approach – one part that studies and characterizes the deterministic structure in the dynamics, and another part that probes for the presence of an underlying stochastic process, or stochastic perturbations of deterministic dynamical systems.

The outline of this paper is as follows. Section 2 provides a mathematical motivation behind using statistical analysis tools to model deterministic chaos. The Perron-Frobenius operator is employed for this purpose. This section also provides motivation behind the potential use of Markov chains for modeling chaotic attractors. Section 3 provides an explanation of how stochastic differential equations with fractal orders can result in strange attractor behavior. Some techniques to quantify this behavior are discussed. In Section 4, some potential future work in this direction is discussed.

2. STATISTICAL ANALYSIS OF DETERMINISTIC CHAOS

One technique for characterizing strange attractors statistically is to learn the probability density function of the attractor's state-space by observing the trajectory over a long period of time. We could, for example, fit the distribution of the observed points on the state-space using a Gaussian Mixture Model (GMM) and learn the parameters using a gradient descent algorithm. Such an approach puts a lower bound on the length of the state-space trajectory required to learn the parameters. An increase in the dimensionality of the state-space and the number of mixtures in the GMM representation will add to this requirement of a large data-size requirement. Another concern with GMM based fitting of attractors is the fact that strange attractors may not have a statistical distribution that is friendly to a GMM fit with a limited number of mixtures.

Another approach to this modeling problem is being studied in the Signal Processing and Communications community [4] [5] [6] . The statistical characteristics of a strange attractor are learned by decomposing the trajectory into a finite number of states, and representing the evolution of the system statistically by a finite Markov chain over these states. Such an approach uses the fact that the trajectory of a dynamical system represents the evolution of the states over a deterministic evolution function. Hence, if the initial conditions were chosen randomly from a certain probability distribution, the evolution of the states resulting from a random sampling of the initial condition will settle down to the same, invariant distribution.

2.1 STATISTICAL DISTRIBUTIONS OF CHAOTIC TRAJECTORIES – THE PERRON-FROBENIUS OPERATOR

For purposes of illustration, consider a one-dimensional map, $M : X \to X$, where X constitutes the state-space of the map. Let us study the statistical behavior of a trajectory generated by such a map. Our goal here is to heuristically motivate on a fundamentally important fact – A highly irregular behavior of chaotic maps in terms of orbit evolutions typically has a remarkably regular statistical structure. We show that this can be extended to sampled flow data as well.

The *Perron-Frobenius operator* (PFO) describes the time evolution of probability densities of the state space under iteration of a deterministic map. This tool is thus useful when we wish to study how a dynamical system which started with initial conditions pulled from a known density function evolves statistically in the state-space.

As an illustration, suppose that the chaotic map M is a non-singular map over a state-space X. The operator $P : L_1 \to L_1$ (where L_1 is the space of Lebesgue integrable functions) characterizes the evolution of the probability density function of the states with each iteration of the map. Using the continuity principle for (measure preserving) deterministic transformation of random variables, it can be shown that:

$$\int_{M^{-1}[0,x]} \phi(\eta) d\eta = \int_X P[\phi(\eta)] d\eta \,, \tag{1}$$

where $\phi(\eta)$ represents the probability distribution of the state space before an iteration of the map was applied to it and $P[\phi(\eta)]$ represents the probability distribution of the state-space after an iteration of the map. Differentiating both sides in equation (1) yields:

$$\begin{aligned} [P\phi]x &= \frac{d}{dx} \int_{M^{-1}[0,x]} \phi(\eta) d\eta \\ &= \int_X \phi(\eta) \delta(\eta - M(x)) d\eta \end{aligned} \,. \tag{2}$$

Note that the Perron-Frobenius operator is a functional, and hence infinite dimensional. In [5] , a finite dimensional approximation to this infinite dimensional operator is presented, by employing a Markov operator of finite rank. This is equivalent to projecting the infinite-dimensional space L_1 (which can be represented by discretely indexed basis functions $\{\xi_i(x)\}_{i=1}^{\infty}$) onto a finite dimensional subspace generated by a subset of the basis functions $\Delta_N = span(\{\xi_i(x)\}_{i=1}^{N})$, such that $\xi_i \in L_1$. The above is equivalent to decomposing the measure space into N (coarse-grained) partitions. Hence, one approximates the probability density of the state space by the finite sum of basis functions

$$\phi(x) \approx \sum_{i=1}^{N} d_i \xi_i(x) \,. \tag{3}$$

In such a representation of the statistics of a discrete dynamical system, Ulam proposed the following matrix:

$$A(i,j) = \frac{m(Y_i \cap M^{-1}(Y_j))}{m(Y_i)} \,, \tag{4}$$

which denotes the relative measure of points in the partition Y_i of the state space that are mapped to points in the partition Y_j under one iteration of the map. With such a finite-Markov approximation in hand, we can use the Perron-Frobenius theorem to estimate the stationary probabilities of the state-space. The Perron-Frobenius theorem guarantees stationary (i.e., independent of the initial state of the system) probabilities for states of a Markov chain, provided the stochastic matrix, $A(i,j)$ for the chain is primitive (i.e., $A^n(i,j) > 0 \ \ \forall \, i,j$, for some finite value of n). For primitive stochastic matrices, the stationary probabilities of the states are given by the normalized Perron vector (the eigen-vector corresponding to a simple eigen-value of one in the spectral decomposition of the matrix). The invariant density in the state space of the map M corresponds to the fixed point of the PFO. This invariant density plays an important role in the estimation of time-averaged statistics of time series generated from nonlinear dynamics.

Figure 1 shows the statistical properties of a state-space partition generated by 1000 iterations of the bend-up bend-down map (section 5). Note that irrespective of whether we generate the initial conditions of the map from a normal distribution or a uniform distribution, the statistical characteristics of the state-space representing the evolution of the trajectory converges to an invariant density. Also note that though the time series generated from two different initial conditions are significantly different (due to

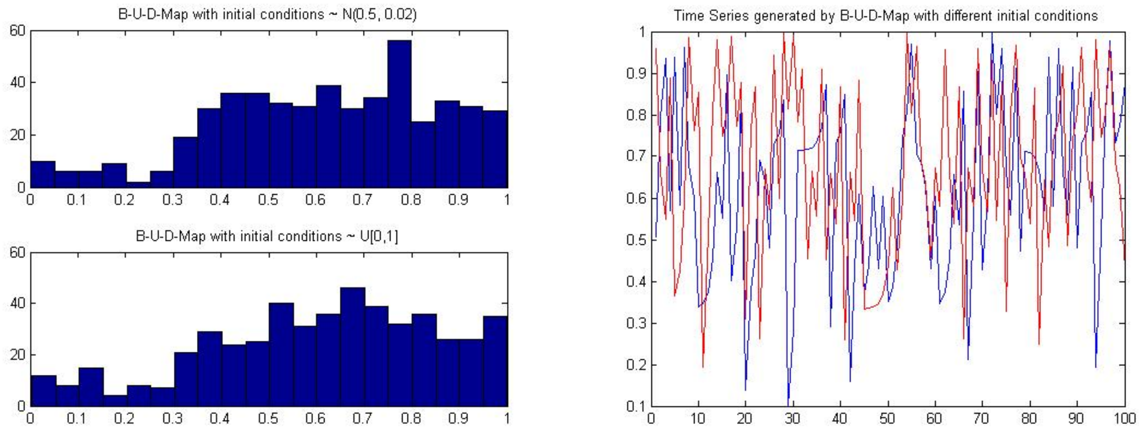

Figure 1 (a) *Statistical distribution of the state space of the Bend-Up, Bend-Down map with initial conditions generated from Normal and Uniform distributions,* (b) *Time series generated by the same map using two different initial conditions.*

the chaotic nature of the map), the statistical properties of the underlying attractor are similar.

The fixed point of the PFO hence provides a means to estimate the statistical distribution of the state-space of the map being studied. If it is desired to characterize the underlying statistical structure of the underlying state-space, one can proceed as follows – (1) Partition the state-space into suitable Markov partitions, (2) Estimate the stochastic matrix for the system over the defined partition, (3) Estimate the long-run (i.e., after transients have died out) statistical structure of the map by using the Perron-Frobenius theorem (which relates the one-step transition matrix of the Markov chain to the invariant distribution over the state-space.)

The above procedure has been successfully applied to synthesis problems in signal processing and communications research [4] [5] [7] .This approach can be extended to an analysis paradigm (e.g. for a pattern classification problem, where one knows the statistical structure of various maps from training data, and wishes to find the distance of an observed time series from the learned models).

As an illustration, consider the following – we have training data (in the form of time series) from two maps (Bend-up, Bend-Down map and the W-map [Appendix]). If we wish to assign a map-label to a test time series, we can obtain the state-sequence of the state-space that resulted in the generation of a time series (e.g., by partitioning the space into Markov partitions and finding the measure of points being mapped from one partition to another in one time step). The problem of identifying the map that generated the test-time series can then be treated in a conventional maximum likelihood framework:

$$\hat{M} = \arg\max_{M} P(s_1 s_2 s_3 ... | M) , \tag{5}$$

which maximizes the probability of the observed state sequence $s_1 s_2 s_3 ...$, given a model (map) M . If we know the stochastic matrix for every map, we can evaluate $P(s_1 s_2 s_3 ... | M)$ using the Markovian assumption of the chain.

Table 1: *Illustrating a confusion matrix using the log-likelihoods of test-time series generated by two maps, compared against reference (trained) models, using a Markov chain representation of the maps.*

	B.U.B.D. -Map	**W - Map**
B.U.B.D. Map	**- 88.1**	**- 110.7**
W-Map	**- 74.5**	**- 49.4**

Table 1 depicts the log-likelihood of test series generated from B.U.B.D. map and W map respectively, computed using the Markov transition probabilities learned from each map. The log-likelihood of the test time series computed using the transition matrix of the correct model is consistently greater than the estimate of the log-likelihood using the wrong model. Further, Figure 2 shows the true statistical distribution of the state space corresponding to two time series, and compares it to the estimate of the distribution using the Perron-Frobenius theory. These results strengthen the claim in this paper – that chaotic maps can indeed be classified using a Markovian assumption.

3. MODELING THE FRACTAL STRUCTURE IN SELF-SIMILAR ATTRACTORS

Many chaotic attractors exhibit self-similarity in their structure. *Self-similarity* is a key concept in fractal geometry, in which, an object appears to be structurally similar at different scales. It is hence natural to define a measure that quantifies this behavior, and use it as a

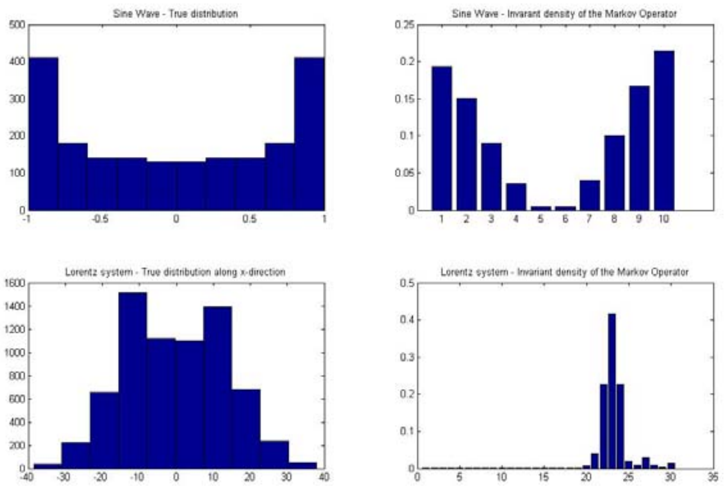

Figure 2: *Statistical distributions of the state space of a sine wave and the x-variable of the Lorentz attractor, and the corresponding invariant density estimate, using the Perron vector of the Stochastic Matrix.*

signature of self-similarity. Figure 3 shows the various types of self-similar fractals. *Deterministic self-similarity* is composed of different features that resemble each other in some way at different scales. Deterministic fractals are typically generated through some Iterated Function System (IFS). *Statistical self-similarity* is composed of features that may change at different length scales, but in such a way that the statistical properties remain same at all resolutions. Statistically self-similar objects are used to model a variety of naturally occurring objects. These can be considered as having being generated by solving stochastic differential equations of fractional orders. The following analysis provides a convincing proof of the fact that fractal objects can indeed be generated as solutions of stochastic differential equations. Consider the following equation:

$$\frac{d^q}{dx^q} f(x) = n(x), \tag{6}$$

where $n(x)$ is white noise. For non-integer values of q, equation (6) has fractional order, since it involves fractional derivatives. Before proceeding with the analysis, let's define fractional derivatives. Although there exist many definitions of fractional derivatives, one definition of interest to us is:

$$\frac{d^q}{dx^q} f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (ik)^q F(k) \exp(ikx) dx, \tag{7}$$

where $F(k)$ is the Fourier transform of $f(x)$. Plugging this back in the original stochastic differential equation,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (ik)^{-q} N(k) \exp(ikx) dk. \tag{8}$$

$f(x)$ can hence be obtained by employing the Liouville-Riemann transform [8] , as:

$$f(x) = \frac{1}{\Gamma(q)} \int_{0}^{x} \frac{n(y)}{(x-y)^{1-q}} dy. \tag{9}$$

It can be shown [8] that equation (9) yields

$$\Pr[f_\lambda(x)] = \frac{1}{\lambda^q} \Pr[f(\lambda x)], \tag{10}$$

which implies scale invariance in the statistics of the signal $f(x)$ - when viewed at a different resolution, the signal will be similar in it's statistical structure. Now that we are convinced that stochastic differential equations with fractal orders leading to (statistically) fractal solutions, let us consider the following analysis / synthesis problems [4]:

1) Forward problem (Equivalent to Synthesis): Given q, determine f .
2) Inverse Problem (Equivalent to Analysis): Given f , determine q .

The inverse problem can be solved if one assumes a certain structure of the characteristic Power-Spectral Density Function (PSDF - $P(k) = |F(k)|^2$). Assuming $P(k) \propto |k|^{-2q}$ (which is a typical spectral representation of fractal signals),

$$P(k) = \frac{c}{|k|^\beta}, \quad \beta = 2q, \tag{11}$$

where c is a constant of proportionality. Consider the case where $P_i \equiv P(k_i)$ is estimated from the FFT of the digital fractal signal. Such a representation can again be approximated by:

$$\hat{P}_i = \frac{c}{|k_i|^\beta}, \quad \beta = 2q. \tag{12}$$

We can use the conventional least-squares formulation to estimate β and c . To convert this into the linear least

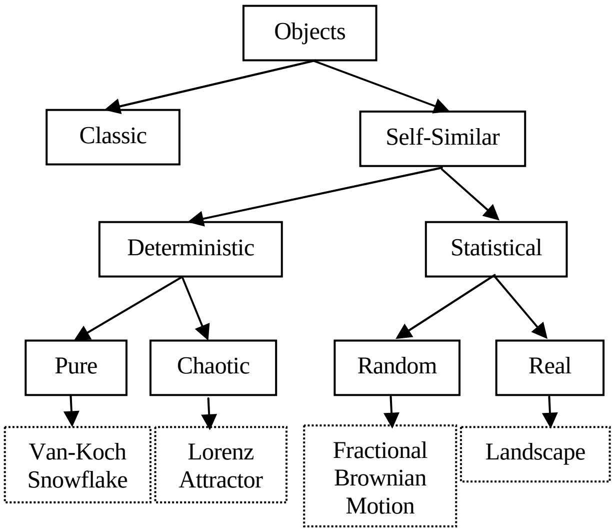

Figure 3: *Different types of Self-Similarities.*

squares setting, let's consider the following error function (on the log scale, instead of the linear scale):

$$\begin{aligned} e &= \sum_i (\ln P_i - \ln \hat{P}_i)^2 \\ &= \sum_i (\ln P_i - [C - \beta \ln |K_i|])^2. \end{aligned} \tag{13}$$

Using the fact that $K_i > 0, P_i > 0\ \forall i$, and solving for β and c such that e is minimized, we obtain:

$$\beta = \frac{N \sum_i (\ln P_i)(\ln K_i) - \left(\sum_i \ln(K_i)\right)\left(\sum_i \ln(P_i)\right)}{\left(\sum_i \ln(K_i)\right)^2 - N \sum_i (\ln(K_i))^2}. \tag{14}$$

In [8] , [9] , the fractal dimension of the attractor is related to the Fourier dimension q as $\beta = 5 - 2D, \beta = 2q$, where β is estimated using equation (14). This hence provides us with a reliable way to estimate the signature of self-similarity in fractal signals.

Many naturally occurring signals have a high frequency decay of the form $P(k) \propto |k|^{-\beta}$, for which the fractal model derived as the solution of a stochastic differential equation is sufficient. Some signals do not have a simple power-law decay of the spectrum. It is hence sometimes desirable to perform a $1/k$ pre-filtering to conform to the requirement of a spectrum decaying with a power law. Alternatively, we need a more general model to incorporate a wider variety of PSDFs. In [8] , the PSDF of the Bermann process is considered.

$$P(k) = \frac{c|k|^g}{k_0^2 + k^2}, \tag{15}$$

for which the corresponding fractal dimension estimate is given by:

$$D = \min\left[\frac{2n}{1-g}, n + \frac{1+g}{2}\right]. \tag{16}$$

4. CONCLUSIONS AND FUTURE WORK

An analysis on the statistical distributions of chaotic maps reveals that they can be well represented by discrete Markov chains. This Markovian property can be extended to multi-dimensional maps, e.g., using a Vector-Quantization approach. Alternatively, the partitioning of the state-space into Markov partitions can be done using a Hidden Markov Model, where the state-space trajectories are treated as observables (emissions) of the hidden states.

Any statistical analysis method that uses a reconstructed phase space (in other words, assumes a purely deterministic system behavior) can not be used to model systems with stochastic inputs. The solution of a stochastic, fractional order differential equation provides a useful tool to quantify the self-similar structure generated by such systems. An important concern that arises when using fractals in signal processing applications is that the analysis of the signal should proceed using the true statistics of the signal, and should not be based on an assumption that the signal obeys some model (e.g., fractal structure). However, if such a model is assumed, and it provides results consistent with the theory, and, it also provides a useful measure for feature extraction and pattern recognition, then such a model-based approach may be desirable. As a note of caution, it must be pointed out that the techniques for modeling / quantifying the fractal structure in an object assume that the underlying structure of the system is indeed of a fractal nature (e.g., a stochastic differential equation of a fractal order). In [9] , the fractal structure in a class of speech sounds has been employed for a satisfactory speech segmentation task. In this work, an analysis window (used to estimate the fractal dimension) is moved by one sample, and regions showing distinct similarity in the fractal structure estimates are segmented.

5. APPENDIX

For completeness, a description of the two chaotic maps used for illustration in this paper is discussed. Some properties of the Perron-Frobenius operator in context to the estimation of the underlying invariant density under chaotic evolution are also discussed.

The bended up-down map is defined as $M : X \to X,\ \ for\ X \in [0,1]$, where,

$$M(x) = \begin{cases} \frac{9x}{2x+3}, & 0 \le x < 1/3 \\ \frac{13x-3}{3x+3}, & 1/3 \le x < 3/5 \\ \frac{29x-15}{7x+3}, & 3/5 \le x < 9/11 \\ \frac{99x-81}{25x-3}, & 9/11 \le x < 1 \end{cases}. \tag{17}$$

The W-map is defined as $M : X \to X,\ \ for\ X \in [0,1]$, where,

$$M(x)=\begin{cases}-4x+1, & 0\le x<1/4\\ 3x-3/4, & 1/4\le x<1/2\\ -3x+9/4, & 1/2\le x<3/4\\ 4x-3, & 3/4\le x<1\end{cases}. \quad (18)$$

Some important properties of the Perron-Frobenius operator (P.F.O.), P are summarized for completeness.

- P is a functional (and hence infinite dimensional) linear operator.
- P is positive, i.e., $P[\phi(\eta)]\ge 0$ if $\phi(\eta)\ge 0$.
- P is measure preserving, i.e., $\int_X [P\phi](x)dx=\int_X [\phi(x)dx$
- The P.F.O. corresponding to the k'th iterate of the map, M^k is $P_k = P^k$.

Another interesting point worth mentioning here is the Markovian behavior of chaotic maps vs. chaotic flow. If we partition the state-space of a chaotic map into Markov partitions and build the one-step transition probability matrix, we are likely to generate a full matrix, since there is no restriction (e.g., continuity) on the evolution of data over a map. However, for flow data (which is obtained as the solution of differential equations), due to the requirement of continuity in the evolution, the corresponding state-transition matrix exhibits a strong block-diagonal structure.

6. REFERENCES

[1] J.P. Eckmann and D. Ruelle, "Ergodic Theory of Chaos and Strange Attractors," *Reviews of Modern Physics*, vol. 57, pp. 617-656, July 1985.

[2] H. Kantz, and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, UK, 2003.

[3] E. Ott, T. Sauer, J. A. Yorke, *Coping with chaos*, Wiley Interscience, New York, New York, USA, 1994.

[4] T. Kohda, "Information Sources Using Chaotic Dynamics," *Proceedings of the IEEE*, vol. 90, pp. 641-661, May 2002.

[5] G. Setti, G. Mazzini, R. Rovatti, S. Callegari, "Statistical Modeling of Discrete-Time Chaotic Processes – Basic Finite Dimensional Tools and Applications," *Proceedings of the IEEE*, vol. 90, pp. 662-690, May 2002.

[6] G. Setti, G. Mazzini, R. Rovatti, S. Callegari, "Statistical Modeling of Discrete-Time Chaotic Processes – Advanced Finite Dimensional Tools and Applications," *Proceedings of the IEEE*, vol. 90, pp. 820-840, May 2002.

[7] S.H. Isabelle, and G.W. Wornell, "Statistical Analysis and Spectral Estimation Techniques for One-Dimensional Chaotic Signals," *IEEE Transactions on Signal Processing*, vol. 45, No. 6, pp. 1495 - 1505, June 1997.

[8] Marwan Al-Akaidi, *Fractal Speech Processing*, Cambridge University Press, UK, 2004.

[9] S. Fekkai, M. Al-Akaidi, J. Blackledge, "Fractal Dimension Segmentation: Isolated Speech Recognition," *Proceedings of the IEE*, 2000.