Detecting Determinism in Speech Phonemes

Xiaolin Liu, Richard J. Povinelli and Michael T. Johnson
Department of Electrical and Computer Engineering
Marquette University, USA

Abstract

This paper[1] presents a discriminating approach to detecting the existence of underlying determinism in speech phonemes using the surrogate data method. The discrimination is made using a statistical measurement of neighboring trajectory directions. This approach is experimentally verified with both deterministic and stochastic time series and then applied to speech phonemes from the TIMIT database. The results show that vowels present some degree of determinism, while no evidence is observed showing that determinism exists with either voiced or unvoiced fricatives.

1. Introduction

Recently suggestions that speech production may be a nonlinear process [1] have produced great interest in applying nonlinear analysis to speech production modeling. A number of studies have been conducted to seek evidence of the possible underlying chaotic features in speech phonemes [2-6]. The most commonly used identification method is to calculate the Lyapunov exponent and correlation dimension, which are invariants of a chaotic system. However, the numerical estimation of Lyapunov exponents can be problematic in the presence of noise [3]. Furthermore, for short data sets, which are generally the case for speech phoneme analysis, the computation of Lyapunov exponents can be difficult. If there are not enough data points in the phase space, increasing the size of the nearest neighborhood can create an erroneous expansion rate. Unfortunately, calculating the correlation dimension, the other invariant, may also be misleading as it is susceptible to false positive results [7]. Thus, studies based on numerically evaluating these two invariants have provided evidence that both support and reject the existence of chaos in speech [2-4].

This paper presents a discriminating approach to detecting the existence of determinism in speech phonemes using the surrogate data method [8]. The implemented discriminating approach is to measure the flow direction of neighboring trajectories within a local subspace in a phase space. The basic notion is

that if trajectories are generated by deterministic equations, then the determinism will cause identical present states (or the neighboring trajectories) to evolve similarly in the near future (the 'analogues' in Lorenz [9]). This similarity can be reflected by detecting the underlying structure of the average difference of the one-step flow direction between neighboring trajectories. The basic idea is that if a time series has some degree of determinism, then the detected structure would be different to that of its corresponding surrogates. However, if a time series comes from a stochastic process, then it is unlikely to present such a phenomenon.

A similar idea of measuring parallel trajectories was first proposed in [10]; however, the underlying algorithm details were not made available. In our approach, improvements to trajectory parallel measure are made in two aspects. Firstly, to locate the nearest trajectory segments to a query point, an algorithm of finding the nearest trajectories is used, instead of using the nearest neighbor algorithm (NNA). Using the NNA can be problematic, because it is highly possible for the NNA to wrongly locate multiple trajectory segments on a single trajectory instead of finding just one nearest segment on each of the nearest trajectories. The paradigm of finding the nearest trajectory is shown in Figure 1. The algorithm of finding the neighboring trajectories was first proposed in [11]. It can be realized by modifying the existing nearest neighbor algorithm. Secondly, in our approach, the importance of choosing a proper embedding delay is highlighted. The accuracy of this approach will be degraded if the time series is over-sampled from a system and the embedding delay is not large enough to show the dynamics [12], because in such a situation, the neighboring trajectories will evolve similarly even for a time series from a stochastic system.
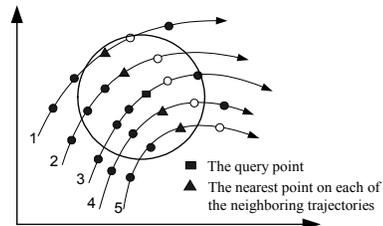
Figure 1. The trajectories 1, 2, 4, and 5 are the four nearest trajectories to the query point. The nearest trajectory algorithm finds only the nearest point, shown as a triangle, on each neighboring trajectory. The rest of dot points on each trajectory are eliminated even though they may be closer to the query point than the triangles.

---

2. Trajectory Direction Measurement (TDM)

The approach to measuring the direction difference between two neighboring trajectories is given in equation (1), which describes the direction difference between two vectors in an arbitrary dimension coordinate system.

$$\cos(\theta) = \frac{a \cdot b}{|a||b|} \qquad (1)$$

in which, $a \cdot b$ is the inner product of vector **a** and **b**, which is given by (2) in a phase space.

$$\begin{aligned} a &= \bar{y}_{n+1} - \bar{y}_n \\ b &= \bar{x}_{n+1} - \bar{x}_n \end{aligned} \qquad (2)$$

$\bar{x}_n$ is the query point and $\bar{y}_n$ is the nearest point on one of the nearest trajectories. $\bar{x}_{n+1}$ and $\bar{y}_{n+1}$ are the next points of $\bar{x}_n$ and $\bar{y}_n$ in the phase space.

The trajectory direction measurement (TDM) is calculated for a variety of embedding dimension. A number of K local subspaces are randomly selected for each embedding dimension. The TDM value of that embedding dimension is calculated by:

$$TDM = \frac{1}{K} \sum_{i=1}^{K} \frac{1}{NT} \frac{1}{\pi} \sum_{j=1}^{NN} \arccos[\frac{a_{ij} . b_i}{|a_{ij}|.|b_i|}] \qquad (3)$$

where b_i is the query trajectory segment of the i_{th} subspace and NT is the number of nearest trajectories selected.

The TDM has values from 0 to 1. A TDM value of 0 means neighboring trajectories in subspaces almost follow the same direction; 0.5 indicates the neighboring trajectories are statistically orthogonal with the query trajectory, and 1 indicates the trajectories in subspaces are evolving towards the opposite direction of the query trajectory.

The calculated TDM value serves as a discriminating approach for the method of surrogate data [8,13]. The method of surrogate data generates an ensemble of surrogate data sets, which mimic the original data set by sharing given properties of the observed time series (such as mean, variance, and power spectrum) while remaining consistent with the null hypothesis. A hierarchy of null hypotheses was proposed in [8].

Of the particular interest the hypothesis we seek to reject is that the signal comes from a linear stochastic process. Under this hypothesis, there are two algorithms of generating surrogate data, Fourier transform (FT) and AR modeling. The FT is selected in our experiment simply because it is a constrained

realization and numerically more stable than AR modeling [8]. This is particularly important when the time series is short because the misestimated coefficients of the AR model will cause the error to diverge when iterating the equation.

3. Experiment with Chaotic and AR Models.

To test the validity of the TDM method, experiments are conducted with time series generated from typical chaotic systems and second-order stochastic systems

3.1 TDM Experiment with Lorenz System
The Lorenz model is defined in (4). The three parameters are configured as typical values.

$$\begin{aligned} dx(t)/dt &= -10\,x(t) + 10\,y(t) \\ dy(t)/dt &= 28\,x(t) - y(t) - x(t)z(t) \\ dz(t)/dt &= -(8/3)z(t) + x(t)y(t) \end{aligned} \qquad (4)$$

A time series of a total 10000 data points is generated with initial values x (1) = 1, y (1) =2, z (1) = 11.005 and $\Delta t = 0.01$ by the fouth-order Runge-Kutta method. The first 2000 data points are discarded to remove transient trajectories. The parameters for the experiments are: 3-10, embedding dimension; 1, time delay; 6, number of surrogate; 4, number of nearest trajectories; 500, number of randomly selected query point.

Figure 2 shows the TDM experiment results with the attractor and power density spectrum of both the original time series x(t) and one of its surrogates. The surrogates have the same power spectrum density with the original time series since the FT is a constrained realization as described in [13].
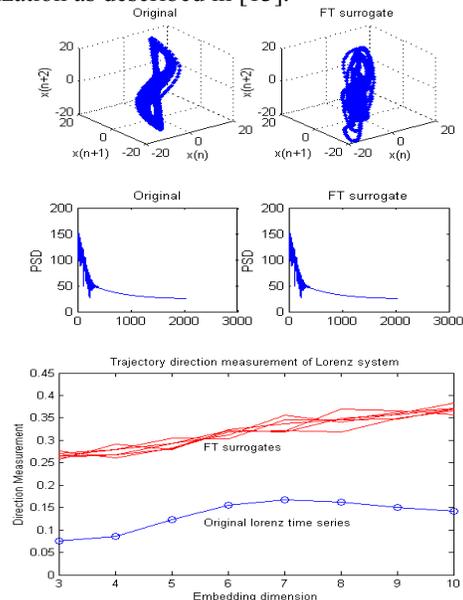
Figure 2. TDM experiment for X variable of Lorenz system

In the Figure 2, the TDM values of six surrogate data sets are clearly distinguishable from the original time series at all embedding dimensions. The underlying structure of one-step flow direction of the original time series does not exist any more with its surrogates. In other words, the information conveyed by determinism in the original time series is destroyed by surrogating the data. Hence, the result clearly suggests that the null hypothesis be rejected.

3.2 TDM Experiment with Logistic Map
Another experiment is conducted with the well-known chaotic map, the logistic map, given by (5).

$$x(t+1) = kx(t)(1 - x(t)) \qquad (5)$$

The logistic map produces a random-like chaotic time series when the coefficient k is greater than 3.57. In this test, k is set to 4. By iterating (5), a time series with 8000 data points is generated. TDM is calculated for embedding dimensions 2 through 10. Other experiment parameters are configured in the same manner as that of the Lorenz test. The TDM experiment results are presented in Figure 3 with the attractor pictures of both original time series and one of the surrogates.

Again in this test, the great difference of the TDM value between the original time series and the 6 surrogate data sets at all embedding dimensions correctly suggests the null hypothesis that the time series comes from a linear stochastic process be rejected.
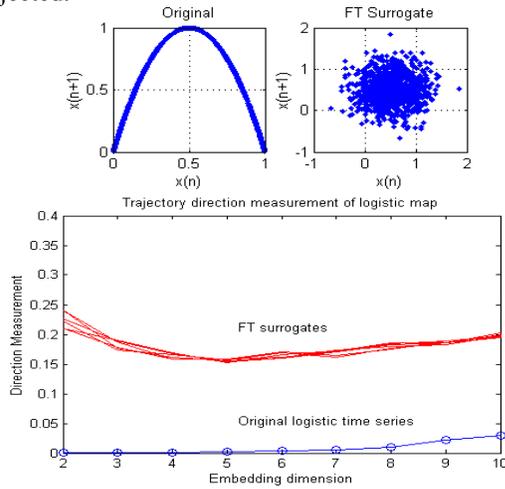
Figure 3. TDM experiments for logistic map

3.3 TDM Experiment with Stochastic Signals
One of our interests is to check whether the TDM method will make a false positive error or not, which means a test rejects the null hypothesis when the null hypothesis is in fact true. Control experiments are conducted with two time series generated from an AR process and a Gaussian white noise generator. The transfer function of AR model is given by:

$$H(z) = \frac{1}{1 + (1/2)z^{-1} + (1/3)z^{-2} + (1/4)z^{-3}} \qquad (6)$$

The experiment parameters are configured as the same with the logistic map test, and the results with Gaussian white noise and AR process are shown in Figure 4 and Figure 5, respectively.

For both cases, the TDM curve of the original time series is totally mixed with that of surrogates at all the embedding dimensions. This again implies that there is no determinism in the original time series that makes the structure of one-step flow direction different to that of its surrogates. Both the original time series and the surrogates share the same characteristic in this aspect. Hence the null hypothesis that the time series comes from a linear stochastic process is accepted correctly.
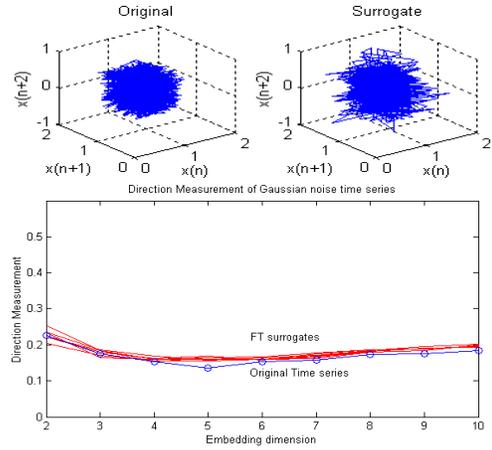
Figure 4. TDM experiments for Gaussian white noise
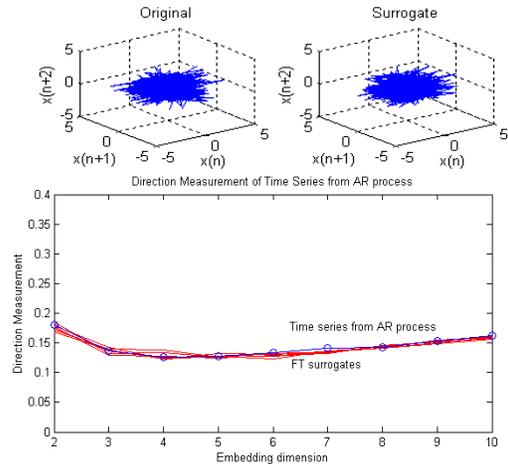
Figure 5. TDM experiments for a time series from an AR model

4. Experiment with Speech Phonemes.

The validity of the TDM method is well founded by the previous experiments. In this section, experiments are conducted to detect the evidence of determinism in English speech phonemes. Voiced and unvoiced phonemes, vowels and fricatives, are randomly selected from the TIMIT database.

To find a proper embedding delay, the average mutual information is calculated for each selected phoneme [12]. Figure 6 shows the result of average mutual information for selected 4 vowels /ao/, /er/, /ow/ and /ay/. The first minimum of the mutual information can be found at about six for these four phonemes.
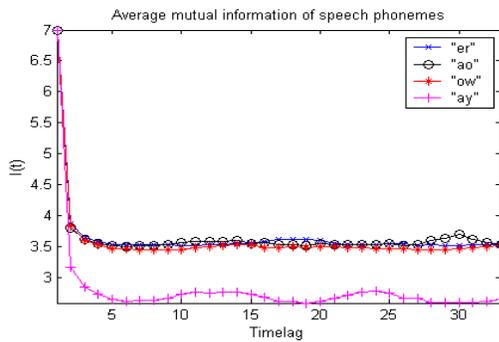
Figure 6. Average mutual information vs. time lag

The FT surrogates of each of the vowels are generated to firstly have a visual inspection of how different they are in the phase space. Interesting differences can be observed from Figure 7 with the original attractors (left column) and one of the surrogates (right column).

The parameters for the experiments are defined as follows: 3-20, embedding dimension; 6, time delay; 12, number of surrogate; 4, number of neighboring trajectories; 400, randomly selected query points. The TDM experiment results are shown in Figure 8.
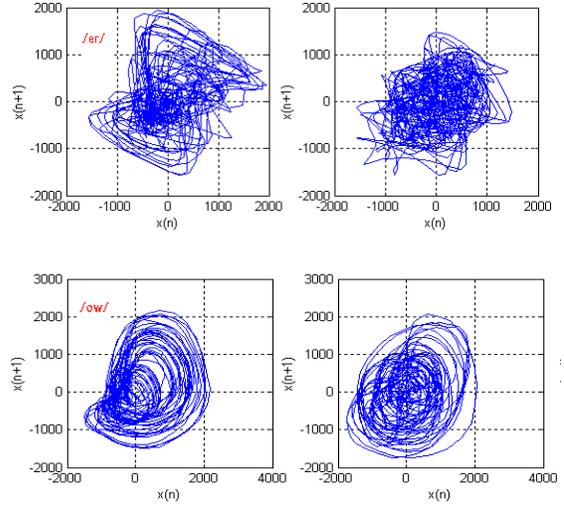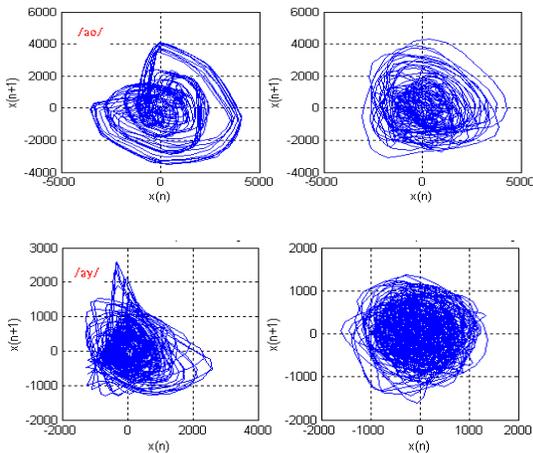
Figure 7. Phase space of four vowels (left) and one of their corresponding surrogate sets (right)
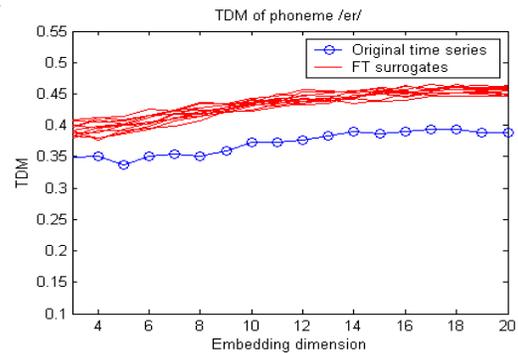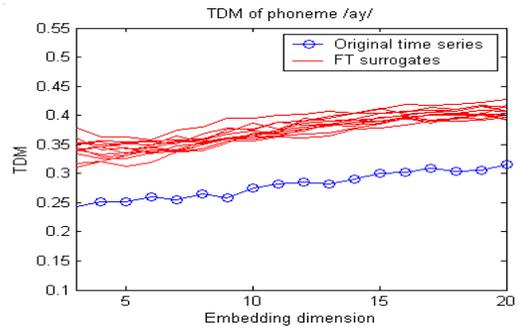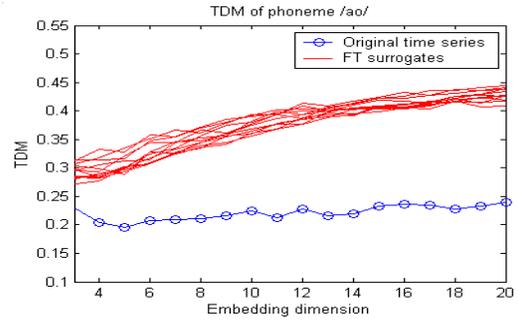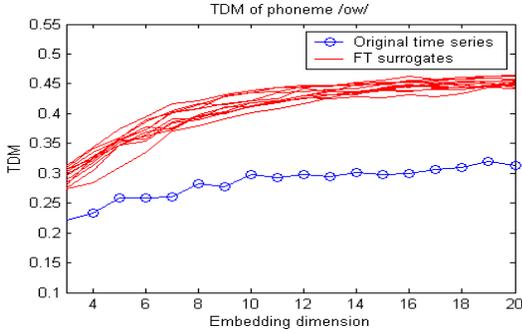
Figure 8. TDM experiments for four vowels

For each of the selected vowels, the curve of the TDM value is clearly distinguishable from its corresponding 12 surrogates. This indicates an existence of determinism in those selected vowels, suggesting a potential application of nonlinear local prediction with vowels for speech processing. Actually this result is also consistent with the fact that each of vowels has a distinct structure in the reconstructed phase space.

Again, the same procedure is conducted for the experiments with one unvoiced fricatives /s/ and one voiced fricative /z/. The experiment results are shown in Figure 9 with the attractors of original one (left column) and one of the surrogates (right column).
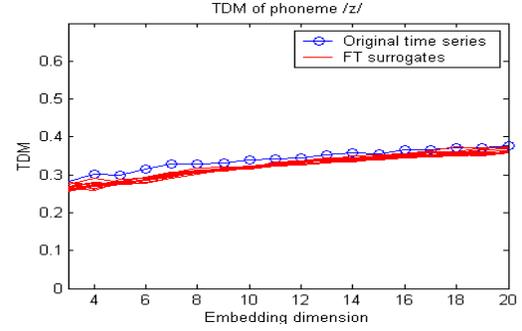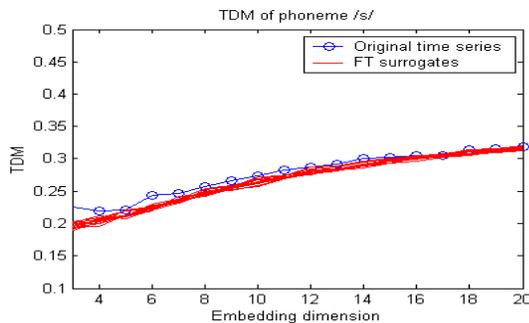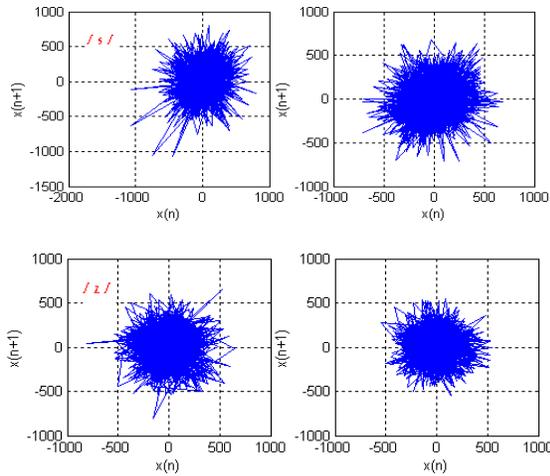
Figure 9. TDM experiments for fricative /s/ and /z/

The mixture of TDM values of the fricative time series and its surrogates indicates that there is no evidence showing an existence of determinism in either voiced or unvoiced fricative example. This suggests that these two fricatives are generated from a stochastic process and no chaos exists in the fricative phoneme production, although it produces a random-like time series.

5. Discussions and Conclusions

This paper presents how to detect determinism in a time series by a statistical direction measurement of neighboring trajectories. Determinism is detected by identifying whether the TDM value of the original time series is significantly different with that of its surrogates. It should be mentioned that evolvement analogue of neighboring trajectories in a deterministic time series actually is a cornerstone of nonlinear local predication [11,14], in which prediction is given by averaging or linearly fitting the coordinate of nearest neighbors. The TDM approach interprets this analogue by a statistical measurement of one-step flow direction of neighboring trajectories. The local linear prediction error can also be used as a discriminating approach for the surrogate data method for determinism detection [8]. However, the TDM approach has the advantage of providing a picture of the flow structure of neighboring trajectories in the phase space.

Reference

[1]   H. M. Teager and S. M. Teager, "Evidence for nonlinear sound production mechanisms in the vocal tract," in *Proc. NATO ASI on Speech Production and Speech Modeling*, pp. 241–261, 1990.

[2]   S. S. Narayanan and A. A. Alwan, "A nonlinear dynamical system analysis of fricative consonants," *J. Acoust. Soc. Amer.*, vol. 97, pp. 2511–2524, Apr. 1995.

[3] M. Banbrook, S. McLaughlin, and I. Mann, "Speech characterization and synthesis by nonlinear methods" *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 1, pp. 1–17, Jan. 1999.

[4] I. Tokuda, R. Tokunaga, and K. Aihara, "A simple geometrical structure underlying speech signals of the Japanese vowel /a/," *Int. J. Bifurc. Chaos*, vol. 6, pp 149-160, 1996.

[5] A. Kumar and S. Mullick, "Nonlinear dynamical analysis of speech," *J. Acoustic Soc. Amer.*, vol. 100, pp. 615-629, 1996.

[6] S. McLaughlin and A. Lowry, "Nonlinear dynamical systems concepts in speech analysis," in EUROSPEECH'93, pp. 377–380, 1993.

[7] P.E. Rapp, "Chaos on the neurosciences: Cautionary tales from the frontier," the *Biologist*, vol. 40, pp. 89-94, 1993.

[8] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer, "Testing for nonlinearity in time series; the method of surrogate data," *Physica D*, vol. 58, pp. 77-94, 1992.

[9] E.N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogues,", *J.Atmos. Sci.*, vol. 26, pp. 636-646, 1969.

[10] Y. Fujimoto and T. Iokibe, "Evaluation of deterministic property of time series by the method of surrogate data and the trajectory parallel measure method," *IEICE trans. Fundamentals of Electronics, Communications and Computer Science*, vol. E83-A, no. 2, pp. 343-349, Feb. 2000.

[11] J. McNames, "A nearest trajectory strategy for time series prediction," *Proc. of the Int. Workshop on Advanced Black-Box Techniques for Nonlinear Modeling*, K.U. Leuven, Belgium, pp.112-128, July 1998.

[12] H. Kantz and T. Schreiber "Nonlinear time series analysis," Cambridge University Press, 1997.

[13] J. Theiler and D. Prichard, "Constrained-realization Monte-Carlo method for hypothesis testing," *Physica D*, vol. 94, pp. 221-235, 1996.

[14] J.D. Farmer and J.J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, pp. 845-848, 1987.