**Optimized nonlinear dynamic analysis of pathological voices with laryngeal paralysis
based on the minimum embedding dimension**

Nanmu Huang

Yu Zhang*

William Calawerts

Jack J. Jiang

Abstract

Objective: The correlation dimension and second order entropy at the minimum embedding dimension were compared to the correlation dimension (D_2) and second order entropy (K_2) based on their efficiency and accuracy in differentiating between normal and pathological voices.

Methods: The minimum embedding dimension was estimated with the Cao method. Nonlinear dynamic parameters, such as correlation dimension and second order entropy, were used to quantitatively analyze the normal and pathologic voice samples.

Results: The computing time of the correlation dimension and second order entropy at the minimum embedding dimension was reduced to approximately one-third of that of traditional D_2 and K_2 calculations, reflecting higher efficiency. The statistical results of linear fitting suggested that the correlation dimension was highly correlated to the correlation dimension at the minimum embedding dimension and second order entropy calculation was highly correlated to the second order entropy at the minimum embedding dimension. Lastly, the results of statistical comparison proved that the correlation dimension at the minimum embedding dimension and second order entropy at the minimum embedding dimension were able to significantly differentiate between normal and disordered voices ($p < 0.001$).

Conclusions: The results suggest that the correlation dimension and second order entropy at the minimum embedding dimension are valid analysis tools for the diagnosis of voice disorders. Additionally, the efficiency and accuracy of these parameters yields potential for clinical usage due to lower computation time than current methods.

Keywords: Minimum embedding dimension; Laryngeal paralysis; Nonlinear dynamic analysis; Correlation dimension; Kolmogorov entropy

1. Introduction

Nonlinear dynamic analysis methods have been proven useful in the study of laryngeal systems.^{1-4, 6-11} Behrman and Baken first investigated the effects of nonstationarity, noise, and finite signal length on the calculation of correlation dimension of electroglottographic signals from normal and pathological voices.¹ Next, Hertrich indicated that there was a significant difference between the fractal dimensions of the electroglottographic signals of subjects with Parkinson's disease and those of normal individuals.² To add to this research, Giovanni found that pathologic voices from patients with unilateral laryngeal paralysis had significantly higher maximal Lyapunov exponents than normal voices.³ Recently, Zhang revealed that both correlation dimension and second-order entropy of pathologic human voices showed a statistically dramatic reduction after surgical excision of vocal polyps, suggesting functional improvements.⁴ The combination of these results demonstrate that nonlinear dynamic analysis methods are effective in statistically analyzing pathologic voices. Although, nonlinear dynamic analysis methods have an extremely large potential for clinical application by accurately diagnosing laryngeal pathologies,^{3, 9, 12, 13} and evaluating therapeutic effects,^{4, 12} they are time consuming in practical applications, especially in cases dealing with large volumes of data. Hence, improving the efficiency of nonlinear dynamic analysis for clinical applications is a crucial issue.

Inspired by the Cao method for its high efficiency in determining the minimum embedding dimension, we investigated the efficacy of correlation dimension and second order entropy at the minimum embedding dimension calculation in our study. Unlike traditional correlation dimension and second order entropy calculations, the correlation dimension and second order entropy at the minimum embedding dimension are not estimated with the increase of embedding dimension d , but estimated merely at the minimum embedding dimension in the scaling region of the radius r , leaving out data of

other embedding dimensions. Thus, the calculations of correlation dimension and second order entropy at the minimum embedding dimension require less input data, reducing computation cost and increasing the speed of the calculation. If these calculations are proven equally as effective as traditional nonlinear analysis methods, they offer a clear advantage.

The purpose of this paper is to make statistical comparisons between the correlation dimension and the correlation dimension at the minimum embedding dimension, denoted as $D_{2,dmin}$ in terms of efficiency and accuracy. The same comparisons will be used between the second order entropy and second order entropy at the minimum embedding dimension, denoted as $K_{2,dmin}$. First, computing time of $D_{2,dmin}$ and $K_{2,dmin}$ and of D_2 and K_2 will be compared. Next, we will assess the abilities of $D_{2,dmin}$ and $K_{2,dmin}$ to differentiate between normal and pathological vocal function. We hypothesize that $D_{2,dmin}$ and $K_{2,dmin}$ will distinguish between normal and pathological voices as effectively as traditional calculations of D_2 and K_2 , while requiring less computation time.

2. Methods

2.1 Subject selection and initial analysis

Sustained vowel recordings from 44 normal subjects and 21 patients with laryngeal paralysis were selected from the Disordered Voice Database and program model 4337, version 1.03 (Kay Elemetrics Corporation, Lincoln Park, NJ), which is developed by the Massachusetts Eye and Ear Infirmary Voice and Speech Lab.⁵

Next, the signals were analyzed using the minimum embedding dimension calculations derived from the Cao method. Then, the nonlinear dynamic methods of D_2 , $D_{2,dmin}$, K_2 , $K_{2,dmin}$ were calculated. Lastly, a Logistic map, Hénon map and Lorenz map were created in order to compare D_2 and $D_{2,dmin}$, and K_2 and $K_{2,dmin}$.

2.2 Nonlinear dynamic analyses of acoustical time series

The dynamics of each voice were reconstructed in a phase space, and then used to calculate

correlation dimension and second-order entropy. For a time series x , sampled at the time interval Δt (Δt is the sampling rate), a phase space can be reconstructed with a time delay vector,

(1)

, where τ is the time delay and d is the embedding dimension.¹⁴ In this paper, an appropriate τ was estimated by using the mutual information method proposed by Fraser and Swinney to find the optimum time lag between coordinates.

(1). *Minimum embedding dimension*

The minimum embedding dimension of each sample is determined as¹⁵:

(2)

where d_i is some measurement of Euclidian distance and is given by the maximum norm, i.e., $d_i = \max_j |x_j - x_{j+\tau}|$; x_i is the i th reconstructed vector with embedding dimension $d + 1$, i is an integer such that i is the nearest neighbor of x_i in the d -dimensional reconstructed phase space in the sense of distance defined above. If two points that stay close in the d -dimensional reconstructed space are still close in the $(d + 1)$ -dimensional reconstructed space, such a pair of points is called true neighbors. Otherwise, they are called false neighbors.

(3)

where $\bar{d}$ is the mean value of all d_i 's and dependent only on the dimension d and the lag τ . It has been found that $\bar{d}$ stops changing when d is greater than some value d_{min} if the time series comes from an attractor. Then d_{min} is the minimum embedding dimension.

Furthermore, Cao defined another quantity which is useful to distinguish deterministic signals from stochastic signals.

(4)

where λ is the Lyapunov exponent. Since the future values are independent of the past values, λ will be equal to 1 for any d for random data. On the contrary, for deterministic data, λ is certainly related to d . As a result, it cannot be a constant

for all d . In other words, there must exist some d 's under which.

(2). Correlation dimension D_2

The correlation dimension D_2 is a quantitative measure that specifies the number of degrees of freedom needed to describe a dynamic system. The correlation dimension can be calculated as^{4, 16, 17}:

$$D_2 = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}, \quad (5)$$

Where r is the radius around x_i , and the correlation integral is

$$C(r) = \frac{1}{N(N-1)} \sum_{i \neq j} H(r - |x_i - x_j|), \quad (6)$$

Where the Heaviside function H satisfies

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (7)$$

With the correlation dimension method, chaos could be distinguished from white noise through its

different dynamic characteristics. The estimated D_2 of white noise does not converge with the increase of embedding dimension d . However, the estimated D_2 of a chaotic system converges to finite value with the increase of d . In addition, a more complex system has a higher dimension, meaning that more degrees of freedom may be needed to describe its dynamic state.

The correlation dimension at minimum embedding dimension can be calculated as:

$$D_2 = \frac{1}{m} \sum_{i=1}^m D_2^i \quad (8)$$

(3). Kolmogorov entropy K_2

Kolmogorov entropy K_2 reveals a loss ratio of information in a dynamic system^{4, 16, 17}, which is also a useful measure to discriminate chaos from stochastic noise. For a regular system and a fixed point, $K_2=0$; for a stochastic system, K_2 approaches infinity; and for a chaotic system, K_2 is a positive constant. That is, the larger Kolmogorov entropy is, the more information lost and the more complex a system is.

Kolomogorov entropy is usually estimated by second order entropy K_2 . The relationship between K_2 and in an d embedding dimensional phase space is as follows:

(9)

Where discrete time interval and embedding dimension d are fixed values.

The second order entropy at minimum embedding dimension is calculated as:

(10)

2.3 Statistical analysis

For all normal and pathological voice samples, D_2 and $D_{2,dmin}$, K_2 and $K_{2,dmin}$ were compared and evaluated by linear fitting in order to show the correlation between the parameters at the minimum embedding dimension. The Mann–Whitney rank sum test was employed using $D_{2,dmin}$ and $K_{2,dmin}$ as dependent variables and the subject groups as independent variables. SigmaStat 3.5 (Systat Software, San Jose, CA) was used for statistical analysis.

3. Results

Figs. 1(a) and (b) show the waveforms and spectra of normal voice and pathologic voice, respectively. It was clear that normal voice was characterized by a nearly periodic waveform and discrete frequency peaks in the spectrum, which belonged to type 1 signal, defined by Titze¹⁸, and shown in Fig. 1(a). However, the abnormal voice was a type 3 signal, for approximately aperiodic waveform and noise-like broadband spectrum, as shown in Fig. 1(b). Fig. 1(c) displays the values $E1(d)$ and $E2(d)$ of voice segments from Fig. 1(a). In Fig. 1(c), both $E1(d)$ and $E2(d)$ attained their saturation value at $d = 6$, consequently, 6 should be the minimum embedding dimension for the time series from that normal voice. In addition, the minimum embedding dimension for the time series from that of irregular voice samples was 9, as shown in Fig. 2(d).

Figs. 2 (a) and (b) show the correlation integral and its slope of normal phonation from Fig. 1(a). The

curves from the top correspond to the embedding dimension $d=1, 2, \dots, 12$, respectively. The correlation dimension was calculated in its scaling region, with a plateau marked in Fig. 2(b). The estimation of second order entropy was calculated in a similar way as calculating correlation dimension. Fig. 2(c) illustrates the trends of the correlation dimension with embedding dimension increase, where curves I, II, and III correspond to random noise, voice from the normal character, and voice from the patient, respectively. With the increase of the embedding dimension d , the D_2 value of random noise exhibited a non-convergent tendency. Contrarily, the final estimated dimensions of the normal and disordered voices converged to 1.623 and 4.185, respectively. $D_{2,dmin}$ of normal and disordered voices were 1.626 and 4.193, labelled on the corresponding graphs. The relationship between the second-order entropy and embedding dimension is reflected in Fig. 2(d), where curves I and II correspond to voice from the normal subjects and pathological patients, respectively.

With the increase of dimension d , the estimates of the second-order entropy K_2 of the normal and disordered voices converge to 0.097 bits/Ts and 0.164 bits/Ts. Likewise, $K_{2,dmin}$ of normal and disordered voices were 0.092 bits/Ts and 0.167 bits/Ts, marked on the corresponding graphs. D_2 , $D_{2,dmin}$, K_2 , and $K_{2,dmin}$ of patients with laryngeal paralysis were higher than those of healthy subjects.

With maximal embedding dimension (m) set to different values, the computing times of D_2 and K_2 , denoted as $t(m)$, were calculated for Logistic map, Lorenz map and a normal voice signal. The computing time of $D_{2,dmin}$ and $K_{2,dmin}$ were denoted as t_{dmin} . The relative values $t(m)/t_{dmin}$ grew gradually as maximal embedding dimension m increased and the relative values $t(12)/t_{dmin}$ were approximately 3, as illustrated in Fig. 3(a). Fig. 3(b) shows the distributions of the computing time of $D_{2,dmin}$ and $K_{2,dmin}$ and of D_2 and K_2 . The computing time of D_2 and K_2 were measured using 10000 data points with maximal embedding dimension m set to 12 for each sample. The computing time of $D_{2,dmin}$ and $K_{2,dmin}$ were measured using 10000 data points for each sample as well. The results of statistical comparative analysis showed that it took average 48.1s to obtain the estimated D_2 and K_2 , while the average computing time for $D_{2,dmin}$ and

$K_{2,dmin}$ was 15.9s. Additionally, in order to verify the connections between D_2 and $D_{2,dmin}$, and K_2 and $K_{2,dmin}$, a linear fitting method was employed for all samples (normal, abnormal phonations and chaotic maps), with the statistical results given in Figs. 3(c) and (d). The slopes of these two fitted curves, shown in Figs. 3(a) and (b), were 0.971 and 1.017, respectively. The squares of coefficient of determination, R-square, were 0.994 and 0.991 for D_2 vs. $D_{2,dmin}$ and K_2 vs. $K_{2,dmin}$, respectively. The goodness of fit for both were close to 1, revealing the strong correlation in the outcomes.

Figs. 4(a) and (b) show the statistical graph of $D_{2,dmin}$ and $K_{2,dmin}$ for the 44 normal and 21 irregular phonation samples. The results of a Mann–Whitney rank sum test performed on the correlation dimensions and second-order entropies at the minimum embedding dimension revealed that the median $D_{2,dmin}$ and $K_{2,dmin}$ values of the 21 dysphonic subjects were significantly higher than those of 44 healthy subjects ($p < 0.001$). This indicates that these parameters were capable of distinguishing irregular phonation from normal phonation. Consequently, it was demonstrated that as the complexity of the signal increased, so did the calculated $D_{2,dmin}$ and $K_{2,dmin}$.

4. Discussion

Nonlinear dynamic methods, such as correlation dimension and Kolmogorov entropy, recently have been suggested as reliable measurements for pathologic voices analysis, showing great value in clinical application^{3, 4, 9}. The primary objective of our study was to investigate a new parameter with quicker computational time than traditional methods. The results confirmed our hypothesis, and showed that the computation time for $D_{2,dmin}$ and $K_{2,dmin}$ was roughly one-third of that for traditional analysis. Furthermore, the study confirmed the ability of $D_{2,dmin}$ and $K_{2,dmin}$ to differentiate between traditional and pathological voice signal samples.

Before estimating correlation dimension and Kolmogorov entropy, we calculated the minimum embedding dimension of voice samples and three classical chaotic systems (Logistic map, Hénon map, Lorenz map) with the Cao method¹⁵, which is considered to be a relatively objective and efficient method

to determine the minimum embedding dimension. As seen in Figs. 1(c) and (d), the minimum embedding dimension for acoustical signal from a normal subject and a pathological subject were 6 and 9, respectively. Additionally, the $D_{2,dmin}$ and $K_{2,dmin}$ of normal and abnormal phonations were significantly different ($p<0.001$), as shown in Figs. 4(a) and (b). From this, we may safely draw the conclusion that the correlation dimension and second order entropy at the minimum embedding dimension are reliable objective methods to quantitatively differentiate between healthy and pathological voice signals.

The estimations of correlation dimension and second order entropy were based on correlation integral and manually calculated in their scaling region with the increase of the embedding dimension d . For example, the value of D_2 converges as the embedding dimension increases in the scaling region, and the final dimension estimated value is obtained by calculating the mean of correlation dimensions at each embedding dimension. Nevertheless the estimated value of $D_{2,dmin}$ can be acquired by simply calculating the correlation dimension at the minimum embedding dimension in the scaling region. The way to estimate K_2 and $K_{2,dmin}$ are same with those of D_2 and $D_{2,dmin}$. The curves of relative values $t(m)/t_{dmin}$ had an increasing trend with the growth of maximal embedding dimension m , meaning that the larger m was, the more time it took to calculate D_2 and K_2 , as seen in Fig. 3(a). Considering that the samples used in our study are low-dimensional chaotic signals, we hypothesize that the two new parameters would show greater efficiency in evaluating samples with higher-dimensional signals. Furthermore, the computing time of $D_{2,dmin}$ and $K_{2,dmin}$ was reduced to approximately one third of that of D_2 and K_2 , as shown in Fig. 3(b). Thus, the estimates of D_2 and K_2 could be replaced by $D_{2,dmin}$ and $K_{2,dmin}$ to dramatically reduce computation time and enhance efficiency.

5. Conclusions

In this paper, correlation dimension, second-order entropy, correlation dimension and second order entropy at the minimum embedding dimension were used to analyze the sustained vowels of normal subjects and patients with laryngeal paralysis. The computing time for $D_{2,dmin}$ and $K_{2,dmin}$ was roughly one-

third of that for D_2 and K_2 , showing timesaving advantage over traditional analysis. We also examined the correlation between D_2 and $D_{2,dmin}$, as well as K_2 and $K_{2,dmin}$. All estimates of D_2 , $D_{2,dmin}$, K_2 , and $K_{2,dmin}$ of disordered phonations were statistically higher than those of healthy phonations, revealing that $D_{2,dmin}$ and $K_{2,dmin}$ are capable of differentiating normal and irregular voices as well as traditional methods of D_2 and K_2 . Our study leads to the conclusion that the corresponding correlation dimension and second order entropy at the minimum embedding dimension could not only analyze pathologic human voices, but also contribute to remarkably decreased computation time, improving its potential value in clinical applications.

Acknowledgments

This work was financially supported in part by the National Science Foundation of China (Grant Nos. 41276040, 41306169, 11174240, and 41422604), the Natural Science Foundation of Fujian Province of China (Grant No. 2012J06010). The project was sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References

- [1] Behrman A, Baken R. Correlation dimension of electroglottographic data from healthy and pathologic subjects. *The Journal of the Acoustical Society of America*. 1997;102(4):2371-9.
- [2] Hertrich I, Lutzenberger W, Spieker S, Ackermann H. Fractal dimension of sustained vowel productions in neurological dysphonias: an acoustic and electroglottographic analysis. *The Journal of the Acoustical Society of America*. 1997;102(1):652-4.
- [3] Giovanni A, Ouaknine M, and Triglia JM. Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis. *Journal of Voice*, 1999. 13(3): p. 341-354.
- [4] Zhang Y, McGilligan C, Zhou L, Vig M, Jiang JJ. Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps. *The Journal of the Acoustical Society of America*. 2004;115(5): 2270-7.
- [5] Massachusetts Eye and Ear Infirmary [CD-ROM]. Voice Disorders Database, Version 1.03. Lincoln Park, NJ: Kay Elemetrics; 1994.
- [6] Titze IR, Baken R, Herzel H. Evidence of chaos in vocal fold vibration, In Titze IR eds. *Vocal Fold*

Physiology: Frontiers in Basic Science. Singular, San Diego; 1993:143-188.

[7] Herzel H, Berry D, Titze IR, Saleh M. Analysis of vocal disorders with methods from nonlinear dynamics. *Journal of Speech, Language, and Hearing Research*. 1994;37(5):1008-19.

[8] Behrman A. Global and local dimensions of vocal dynamics. *The Journal of the Acoustical Society of America*. 1999;105:432-43.

[9] Jiang J, Zhang Y. Nonlinear dynamic analysis of speech from pathological subjects. *Electronics Letters*. 2002;38(6):294-5.

[10] Berry DA, Herzel H, Titze IR, Krischer K. Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *The Journal of the Acoustical Society of America*. 1994;95(6):3595-604.

[11] Jiang JJ, Zhang Y, Stern J. Modeling of chaotic vibrations in symmetric vocal folds. *The Journal of the Acoustical Society of America*. 2001;110(4):2120-8.

[12] Zhang Y, Jiang JJ, Wallace SM, Zhou L. Comparison of nonlinear dynamic methods and perturbation methods for voice analysis. *The Journal of the Acoustical Society of America*. 2005;118(4):2551-60.

[13] Zhang Y, Shao J, Krausert CR, Zhang S, Jiang JJ. High-speed image analysis reveals chaotic vibratory behaviors of pathological vocal folds. *Chaos, Solitons & Fractals*. 2011;44(1):169-77.

[14] Packard NH, Crutchfield JP, Farmer JD, Shaw RS. Geometry from a time series. *Physical review letters*. 1980;45(9):712.

[15] Cao L. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*. 1997;110(1):43-50.

[16] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Physica D Nonlinear Phenomena*. 1983;9(s 1-2):189-208.

[17] Kantz H, Schreiber T. *Nonlinear time series analysis*. 2nd ed. Cambridge: Cambridge University Press; 2003. p. 75–102 & 212–229.(14)

[18] Titze I. Summary statement: Workshop on acoustic voice analysis. USA: Denver, CO: National Center for Voice and Speech. 1995.

Figure Captions

Figure. 1 (a) The time series (the left curve) and frequency spectrum (the lower curve) of the voice from a normal individual. (b) The time series (the left curve) and frequency spectrum (the lower curve) of the voice from a patient. (c) The values E1 and E2 for the voice data from Fig. 1(a), with time delay = 15. (d) The values E1 and E2 for the voice data from Fig. 1(b), with time delay = 6.

Figure. 2 (a) Correlation integral of the normal phonation, in which the curves from the top correspond to the embedding dimension $d=1,2,\dots,12$, respectively. (b) The slope of Fig. 2(a) vs. $\ln r$. (c) The estimated dimension versus m for random noise, the voice from normal character, and from patient. (d) The estimated entropy versus m for the voice from normal character, and from patient.

Figure. 3 (a) The effects of maximal embedding dimension on computing time for Logistic map, Lorenz map and a normal voice signal. (b) The distribution of computing time of $D_{2,dmin}$ and $K_{2,dmin}$ and of D_2 and K_2 for all normal and abnormal phonations and chaotic maps, where the line inside the box marks the median and the dots are the outlying points. (c) The linear fitting curve of D_2 versus $D_{2,dmin}$ for normal,

abnormal phonations, and chaotic maps. (d) The linear fitting curve of K_2 versus $K_{2,dmin}$ for normal, abnormal phonations, and chaotic maps.

Figure. 4 (a) The distribution of $D_{2,dmin}$ for all normal and abnormal phonations. (b) The distribution of $K_{2,dmin}$ for all normal and abnormal phonations, where the line inside the box marks the median and the dots are the outlying points.