Optimized Nonlinear Dynamic Analysis of Pathologic Voices With Laryngeal Paralysis Based on the Minimum Embedding Dimension

*Nanmu Huang, †,‡Yu Zhang, §William Calawerts, and §Jack J. Jiang, *†Xiamen and ‡Beijing, China, and §Madison, Wisconsin

Summary: Objective. The present study aims to compare the correlation dimension and second-order entropy at the minimum embedding dimension with the correlation dimension (D_2) and second-order entropy (K_2) based on their efficiency and accuracy in differentiating between normal and pathologic 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 yield potential for clinical usage because of lower computation time than current methods.

Key Words: Minimum embedding dimension–Laryngeal paralysis–Nonlinear dynamic analysis–Correlation dimension–Kolmogorov entropy.

INTRODUCTION

Nonlinear dynamic analysis methods have been proven useful in the study of laryngeal systems.[1–10] 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 pathologic voices.[1] Next, Hertrich et al 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 et al found that pathologic voices from patients with unilateral laryngeal paralysis had significantly higher maximal Lyapunov exponents than normal voices.[3] Recently, Zhang et al 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 demonstrates 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,[3,8,11,12] and evaluating therapeutic effects,[4,11] 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 the present study 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

Accepted for publication July 25, 2016.
From the *Department of Applied Marine Physics and Engineering, College of Ocean and Earth Sciences, Xiamen University, Xiamen, Fujian, China; †Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Ministry of Education, Xiamen University, Xiamen, China; ‡State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing, China; and the §Division of Otolaryngology – Head and Neck Surgery, Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.

Address correspondence and reprint requests to Yu Zhang, Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Ministry of Education, Xiamen University, Xiamen 361005, China. E-mail: yuzhang@xmu.edu.cn
Journal of Voice, Vol. 31, No. 2, pp. 249.e1–249.e7
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.07.021

$D_{2,dmin}$ and $K_{2,dmin}$ to differentiate between normal and pathologic vocal functions. We hypothesize that $D_{2,dmin}$ and $K_{2,dmin}$ will distinguish between normal and pathologic voices as effectively as traditional calculations of D_2 and K_2, while requiring less computation time.

METHODS

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.[13]

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

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 $x(t_i) \in \boldsymbol{R}, t_i = t_0 + i\Delta t (i = 1, 2, \ldots, N)$, sampled at the time interval $\Delta t = 1/f_s$ (f_s is the sampling rate), a phase space can be reconstructed with a time delay vector,

$$y_i(d) = (x_i, x_{i+\tau}, \ldots x_{i+(d-1)\tau}), \quad (1)$$

$i = 1, 2, \ldots, N-(d-1)\tau$, 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.

Minimum embedding dimension

The minimum embedding dimension of each sample is determined as:[15]

$$a(i,d) = \frac{\|y_i(d+1) - y_{n(i,d)}(d+1)\|}{\|y_i(d) - y_{n(i,d)}(d)\|}, i = 1, 2, \ldots, N - d\tau, \quad (2)$$

where $\|\cdot\|$ is some measurement of Euclidian distance and is given by the maximum norm, that is, $\|y_k(m) - y_l(m)\| = \max\limits_{0 \le j \le m-1} |x_{k+j\tau} - x_{l+j\tau}|$; $y_i(d+1)$ is the ith reconstructed vector with embedding dimension $d+1, n(i,d) (1 \le n(i,d) \le N - d\tau)$ is an integer such that $y_{n(i,d)}(d)$ is the nearest neighbor of $y_i(d)$ in the d-dimensional reconstructed phase space in the sense of distance $\|\cdot\|$ 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.

$$E1(d) = E(d+1)/E(d), \quad (3)$$

where $E(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} a(i,d)$ which is the mean value of all $a(i,d)$'s and dependent only on the dimension d and the lag τ. It has been found that $E1(d)$ stops changing when d is greater than some value d_0 if the time series comes from an attractor. Then $d_0 + 1$ is the minimum embedding dimension.

Furthermore, Cao defined another quantity $E2(d)$, which is useful to distinguish deterministic signals from stochastic signals.

$$E2(d) = E^*(d+1)/E^*(d), \quad (4)$$

where $E^*(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} |x_{i+d\tau} - x_{n(i,d)+d\tau}|$. Because the future values are independent of the past values, $E2(d)$ will be equal to 1 for any d for random data. On the contrary, for deterministic data, $E2(d)$ 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 $E2(d) \neq 1$.

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]

$$D_2(d) = \lim_{r \to 0} \lim_{N \to \infty} \frac{\ln C(N,d,r)}{\ln r}, \quad (5)$$

where r is the radius around $y_i(d)$, and the correlation integral $C(N,d,r)$ is:

$$C(N,d,r) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \theta(r - \|y_i(d) - y_j(d)\|), \quad (6)$$

where the Heaviside function $\theta(x)$ satisfies:

$$\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0 \end{cases}. \quad (7)$$

With the correlation dimension method, chaos could be distinguished from white noise through chaos's 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:

$$D_{2,dmin} = D_2(d)|_{d=dmin}. \quad (8)$$

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 is lost and the more complex a system is.