# Using Fractal Dimension Analysis on Objective Auscultation of Traditional Chinese Medical Diagnosis

Chuang-Chien Chiu* Ming-Ta Yang Chern-Sheng Lin

*Institute of Automatic Control Engineering, Feng Chia University, Taichung, Taiwan, 407, ROC*

Received 17 October 2002; Accepted 8 November 2002

## Abstract

Speech variation is important for clinical examinations in auscultation of traditional Chinese medicine. This study applied continuous speech to compare the difference among sixteen qi-vacuity patients, nine ying-vacuity patients and twelve normal subjects. In order to compute the speech characteristics related to production irregularities, the fractal dimension (FD) parameter was used. Two sets of test sentences were spoken in Mandarin: one is specific design and the other is random choice. Each test sentence was pronounced three times as the test patterns, and then we compared the difference of complexity among them using the dynamic time warping (DTW) technique. The results revealed: (1) although the same sentence was read by each subject three times, the distributions of FD values each time were not exactly the same to each other; (2) no significant correlation was found between speech variations and syllables; (3) the utterances of test sentences of both qi-vacuity and ying-vacuity patients have much more variations than that of normal subjects; (4) good classification results (over 85%) were obtained using both sets of test sentences for the classification between deficient and normal subjects; (5) the classification accuracy using FD values of continuous speech was better than the results using five acoustic parameters of the vowel /a/ in our previous study. In conclusion, the fractal dimension analysis of continuous speech provides an alternative and outstanding approach to the objective auscultation of Chinese medical diagnosis.

***Keywords*:** Auscultation, Traditional Chinese medicine, Fractal dimension, Dynamic time warping

## Introduction

Speech is not only a tool for humans to express emotion and communication, but also to reflect the condition of their health. Speech variation is important for clinical examinations in auscultation of traditional Chinese medicine (TCM). However, the auscultation process has been anecdotal and rarely discussed due to the lack of objective assessment and accurate quantification [1]. Even in western medicine, an objective evaluation of the laryngeal pathology by vocal changes has been difficult for the laryngologist or speech pathologist to quantify [2]. One successful example has been performed based on the acoustic analysis of pathological speech qualitatively and quantitatively. Accardo and Mumolo [3] applied traditional acoustic pattern matching algorithms to discriminate between normal and pathological speech. In their study, the pathological speech has been obtained from subjects with Friedreich's ataxia, a degenerative disease of the central nervous system which is characterized by speech rate impairments, disorders of the articulatory movements, hypotonia and tremor. For each utterance, the short-time fractal dimension (FD) parameter and, for comparison, the zero-crossing and energy ratio parameters were evaluated and used in the classification task by means of a dynamic programming procedure. They concluded that the FD parameter seemed to provide a more reliable pattern classification than zero-crossing and energy ratio.

On the other hand, qualitative evaluation by physicians for auscultation in TCM based on vocal changes for determining the nature and location of the disease is still very difficult to quantify objectively. Not all physicians have the same sense of voice, which in turn, influences their subjective judgments. A quantitative analysis should produce the possibility of evaluating and classifying speech in a less subjective way and derive an automatic diagnostic procedure. With the great progress of science, we can obtain more useful information from physiological signals via speech analysis methods. In this way, auscultation of traditional Chinese medicine would attain both qualitative and quantitative standards necessary for future applications in clinical environment or medical education.

In our previous study [1], we aimed at applying the modern techniques of voice analysis to analyze and identify the characteristics of the non-vacuity, qi-vacuity, and ying-vacuity in subjects. The vowel /a/ was chosen as the utterance for each patient to pronounce in that research.

* Corresponding author: Chuang-Chien Chiu
Tel: +886-4-24517250 ext.3930; Fax: +886-4-24515701
E-mail: chiuc@auto.fcu.edu.tw

According to the theory of Chinese medicine, states of qi, ying, yang and blood changing with the viscera function may lead to voice change. For instance, a qi-vacuity patient has a faint voice [4]. Indeed, the invasion of external microbes or functional disorders into the body exerting an effect on the voice has been widely accepted [5]. In a physiological sense [6], qi-vacuity means weakness of the qi. This is usually associated with poor organic functions and a general weakening from illness or overwork, dietary imbalance, or damage to the ying-blood. Ying-vacuity is the manifestation of insufficiency of the ying aspect and depletion of body fluids and blood. Yang-vacuity is a manifestation of an insufficiency of yang-qi with a reduction in the warming and activating power of the body. A pale white or withered-yellow facial complexion, dizziness, relatively pale tongue, and a weak pulse characterize blood-vacuity. Four novel acoustic parameters, the average number of zero-crossings (A1), the variations in local peaks and valleys (A2), the variations in first and second formant frequencies (A3), and the spectral energy ratio in high frequency (3,000 Hz) (A4), were presented to analyze and identify the characteristics among non-vacuity, qi-vacuity, and ying-vacuity subjects. Among these acoustic parameters, two temporal parameters, A1 and A2, outperformed other parameters in classifying both non-vacuity and deficient subjects. A4 was adequate for the classification between qi-vacuity and ying-vacuity patients [1].

There are many studies demonstrating this subject [7-10]. In most of their studies, only vowel was used as the utterance in their experiments. However, speech includes not only vowels but also consonants. The auscultations contain both vowels and consonants were more commonly used to the auscultation of traditional Chinese medical diagnosis than vowels only. The impetus of this study is to use the pattern matching method based on fractal dimension algorithm for continuous speech, proposed by Accardo and Mumolo [3], to identify the difference of the auscultation characteristics among normal persons, qi-vacuity and ying-vacuity patients. This work represents an early step in an ongoing effort to develop an automatic method to discriminate the continuous speech between normal and deficient subjects for objective auscultation of TCM.

## Material and methods

### *Subjects*

Twelve subjects (7 males and 5 females, mean age 38) formed the control group. The deficient groups of subjects were recruited from the outpatient department of internal medicine of Chinese Medicine Department, China Medical College Hospital, Taiwan R.O.C. All subjects were chosen by the criteria of standard vacuity of diagnosis [11]. None of the subjects had any voice or throat complaints at the time of these experiments. Sixteen subjects (7 males and 9 females, mean age 44) formed the qi-vacuity group. Nine subjects (8 males and 1 female, mean age 37) formed the ying-vacuity group. The information of the groups, sex and ages of all subjects in the experiments is listed in Table 1.

Table 1. The groups, sex and ages of all subjects in the experiments.

| | Subjects | | | Age | | |
|---|---|---|---|---|---|---|
| | Male | Female | Total | Max. | Min. | Mean±SD |
| Normal | 7 | 5 | 12 | 60 | 24 | 38 ± 17.1 |
| Qi-vacuity | 7 | 9 | 16 | 71 | 23 | 44 ± 13.6 |
| Ying-vacuity | 8 | 1 | 9 | 71 | 12 | 37 ±21.2 |

### *Speech Data*

The test sentences were designed to meet the following three requirements:

(1) The sentences were shallow, clear and easy to understand.

(2) Each word in the test sentences should contain both vowels and consonants simultaneously.

(3) The sentences should be read fluently without any difficulty.

By these criteria, we designed the test sentence ($S_1$) and random choice sentence ($S_2$) spoken in Mandarin (details information are shown in [12]).

We considered difference in personal nonstandard accents and the tone or manner of speaking, etc. In this situation, we compared the pronounced variation at different times with the same person. The test sentences were pronounced three times as the test patterns in a short-term interval. Then we observed the variations between them.

The recording environment has good air condition (the temperature was controlled in 24±2°C). At first, we allowed patients to rest for five minutes and then play the standard speech sound and told them try to imitate it. And then told the patients to read the sentences three times. The speech signals were digitized using a 16-bit A/D converter (Model AT-DSP2200 manufactured by National Instruments) at 10 KHz sampling rate with anti-aliasing function. To distinguish the variations between three repetitions of the speech data, the same test sentence should be read three times with an interval time of six seconds in between each repetition. In addition, to verify and relate results with the previously study, the vowel /a/ was also recorded.

### *Speech analysis*

In order to quantify the speech characteristics related to production irregularities, the FD parameter was used. The irregularities can be considered produced by noise, i.e. by random fluctuation or stochastic processes, but they could also be due to chaos, i.e. to irregularities that arise in a non-linear deterministic system [3]. We used the box counting method to evaluate the FD of each sentence. This is a convenient strategy allowing us to quantify the complexity of the signal in the temporal succession of events [13-16]. The FD parameter, $D$, is defined as follows [17]:

$$D = \lim_{n \to \infty} \{ \frac{\ln(N_n(A))}{\ln(2^n)} \} \qquad (1)$$

Where $A$ is the sequence of the speech feature, $A = a_1, a_2 \ldots a_n$, $a_n$'s are sampling points. $N_n(A)$ is the number of boxes filled by events as the resolution at the $(1/2)^n$. $D$ is the FD parameter. The FD algorithm and calculation procedure are

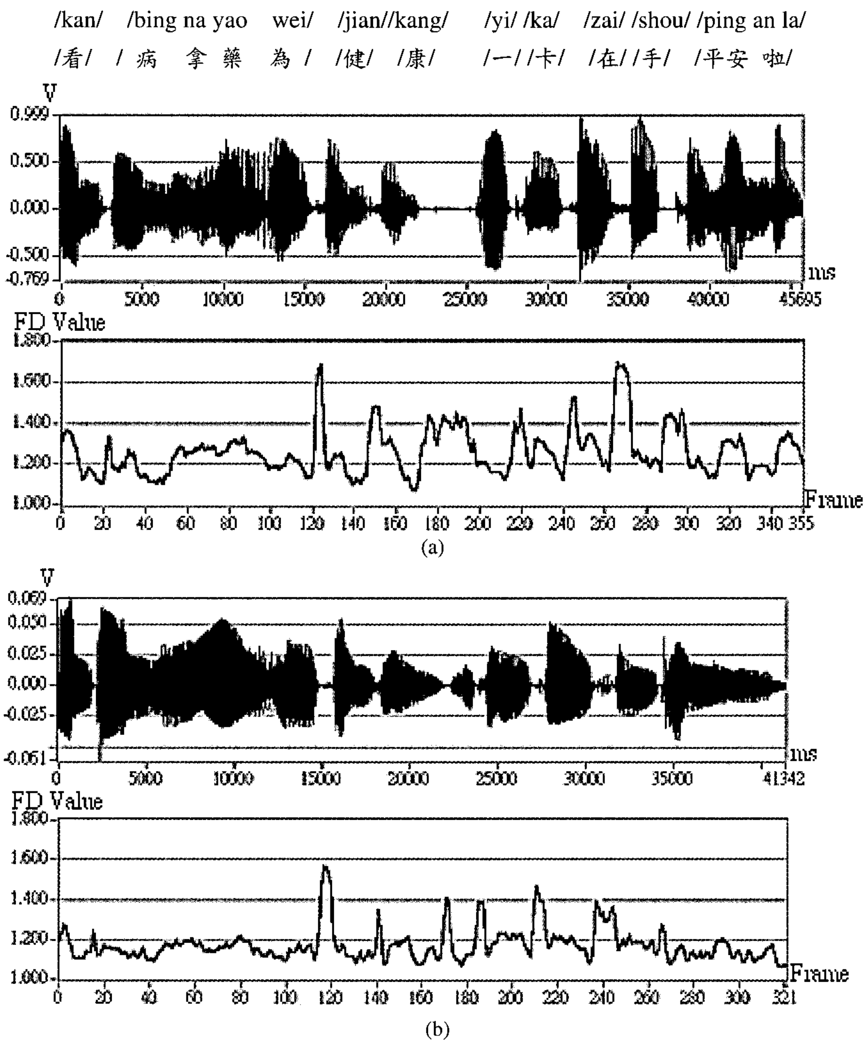


Figure 1. Typical speech waveforms and their corresponding FD parameter values for (a) a normal subject, (b) a qi-vacuity patient. Note that the test sentence (/kan//bing na yao wei//jian//kang/ /yi//ka//zai//shou//ping an la/) is spoken.

listed as follows.

(1) Assume the box counting must approximate the smallest number of squares $A$ necessary to cover the speech feature.

(2) In the continuity of the functions, we retain the highest and lowest value in each column. Every box in between will be entered by the function.

(3) Starting with a number of samples that is a power of two, we halved the number of columns recursively.

(4) Using the results of (3), we plot $\ln(N_n(A))$ *vs.* $\ln(1/2)^n$ points in log-log graph, and then find the approximate straight line in these points.

(5) Fractal dimension $D$ is obtained by the slope of the straight line of (4)

The speech signal for FD analysis purpose was divided into 256 ms for each speech frame with a 32 ms overlapping between them. In Figure 1, typical speech waveforms and their corresponding FD parameter values for a normal subject and a qi-vacuity patient with the test sentence ($S_1$) being spoken are shown. Also, typical speech waveforms and their corresponding FD parameter values for a normal subject and a qi-vacuity patient with the test sentence ($S_2$) being spoken are shown in Figure 2.

Table 2. The distance evaluation results using DTW algorithm of the example shown in Figure 1: the intra-normal and intra-qi-vacuity distance evaluation of $D$ variations.

| Distance ($E$) | First vs. Second | First vs. Third | Second vs. Third | Total |
|---|---|---|---|---|
| Normal | 6.760 | 4.877 | 5.296 | 16.933 |
| Qi-vacuity | 7.792 | 6.697 | 6.432 | 20.921 |

Table 3. The distance evaluation results using DTW algorithm of the example shown in Figure 2: the intra-normal and intra-qi-vacuity distance evaluation of D variations.

| Distance ($E$) | First vs. Second | First vs. Third | Second vs. Third | Total |
|---|---|---|---|---|
| Normal | 6.726 | 4.480 | 4.065 | 15.271 |
| Qi-vacuity | 9.485 | 8.651 | 4.491 | 22.627 |

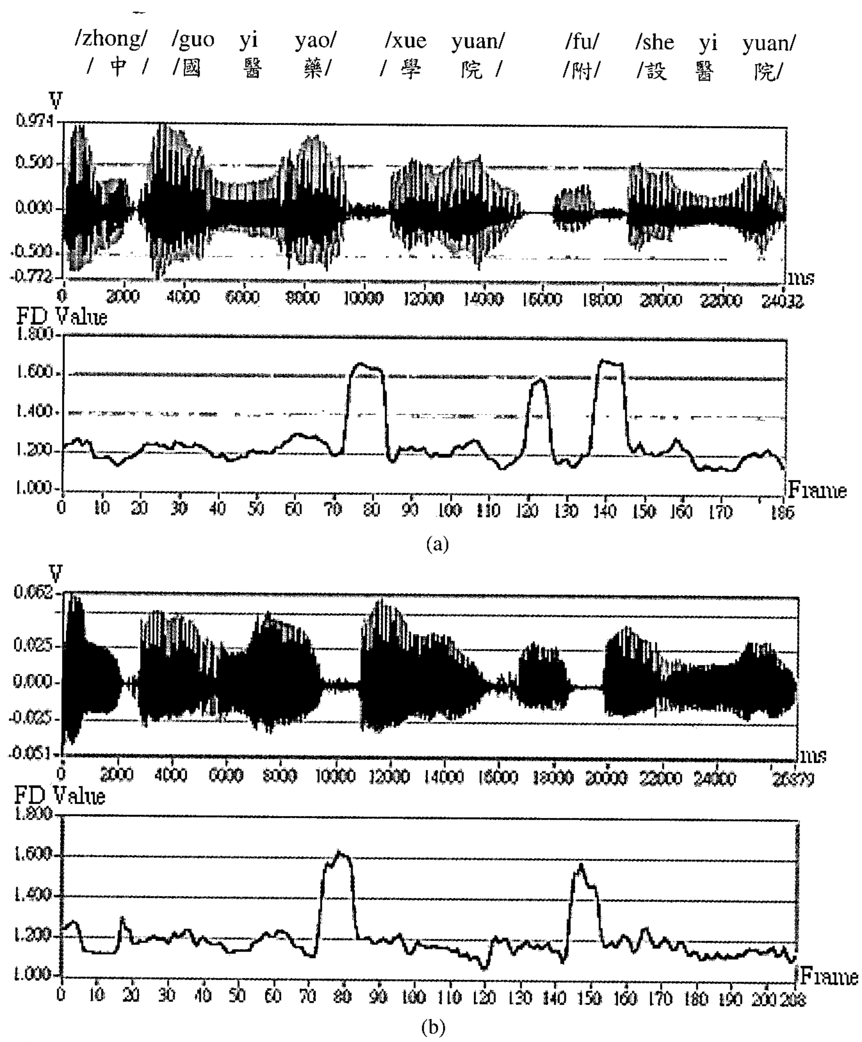


Figure 2. Typical speech waveforms and their corresponding FD parameter values for (a) a normal subject, (b) a qi-vacuity patient. Note that the random choice sentence (/zhong//guo yi yao/ /xue yuan/ /fu/ /she yi yuan/) is spoken.

Table 4. The statistical analysis results among normal, qi-vacuity, and ying-vacuity groups in two sentences.

| Sentence | Group | Mean±SD | *p* value |
|---|---|---|---|
| $S_1$ | Normal | 15.158±2.471 | $p<0.05$* |
| | Qi-vacuity | 23.448±8.064 | |
| | Normal | 15.158±2.471 | $p<0.05$* |
| | Ying-vacuity | 24.322±7.167 | |
| | Qi-vacuity | 23.448±8.064 | $p<0.05$* |
| | Ying-vacuity | 24.322±7.167 | |
| $S_2$ | Normal | 14.257 ±2.346 | $p<0.01$* |
| | Qi-vacuity | 25.354 ±4.915 | |
| | Normal | 14.257 ±2.346 | $p<0.01$* |
| | Ying-vacuity | 34.677±7.859 | |
| | Qi-vacuity | 25.354 ±4.915 | $p<0.01$* |
| | Ying-vacuity | 34.677±7.859 | |

* represents that there exists statistical significance between two groups of subjects.

*Procedures*

When the *D* variations of the speech were computed, we compared the differences of each pair of sentences from among the three utterances. A possible problem with this technique is how to compute the differences in the path length of fractal dimension values. The dynamic time warping (DTW) method was used to compare the differences of these path distributions of fractal dimension [18-19]. The values of *D* variations were then normalized for saving computer executions time. The distance between two lines was calculated in Euclidean distance [20]. Finally, the optimal distance *E* was calculated by DTW algorithm and it can be considered as the difference intensity between two *D* distributions. For instance, the distance evaluation results using DTW algorithm of the examples shown in Figure 1 and Figure 2 were depicted in Table 2 and Table 3 respectively.

## Results

In order to test the performance of the proposed features, the Student's t-test analysis in the SPSS statistical analysis package was implemented. The values of *E* in each group of subjects were calculated. All parametric data were expressed

Table 5. The t-test results between normal and deficient groups in two sentences.

| Sentence | Group | Mean±SD | *p* value |
|---|---|---|---|
| $S_1$ | Normal | 14.257±2.346 | $p<0.05$* |
| | Deficient | 22.393±4.970 | |
| $S_2$ | Normal | 14.257±2.346 | $p<0.01$* |
| | Deficient | 29.217±6.769 | |

* represents that there exists statistical significance between two groups of subjects.

Table 6. The accuracy of discrimination among normal, qi-vacuity, and ying-vacuity groups.

| | $S_1$ | $S_2$ | Merge two sentences |
|---|---|---|---|
| Normal vs. Qi-vacuity | 81.2% | 92.9% | 85.7% |
| Normal vs.Ying-vacuity | 72.0% | 90.5% | 85.7% |
| Qi-vacuity vs. Ying-vacuity | 71.4% | 80.0% | 80.0% |

Table 7. The accuracy of discrimination between normal and deficient groups.

| | $S_1$ | $S_2$ | Merge two sentences |
|---|---|---|---|
| Normal vs. Deficient | 78.4% | 83.8% | 86.5% |

Table 8. The accuracy of discrimination between normal and ying-vacuity groups.

| Voice parameters | A1 | {A1, A2} | {A2, A3} | A2 | A3 |
|---|---|---|---|---|---|
| Accuracy | 75.0% | 71.4% | 71.4% | 64.3% | 64.3% |

as mean±SD. The two-sided paired t-test was applied to test between two groups of subjects. The statistical significance was defined as $p<0.05$. The results are presented in Table 4.

In order to evaluate the normal and deficient groups, the qi-vacuity and ying-vacuity subjects were combined as the deficient group. We used the t-test to evaluate the statistical significance of difference between the normal and deficient groups. The results are shown in Table 5.

In comparison with the accuracy between physician's diagnosis and the discriminant results, shown in Table 6 and Table 7, we suggested that the Fisher linear discriminant function could be used as a method in auscultation diagnosis.

Also, the previous methods [1] were applied to classify the groups between normal and deficient subjects. The vowel /a/ was estimated by the proposed acoustic parameters, and discriminated the accuracy among normal, qi-vacuity, and ying-vacuity groups. The results were showed in Table 8~ Table 11. For simplicity, we only show results with better discriminant performance. Note that a new acoustic parameter, the spectral energy ratio in low frequency (800 Hz) (A5), was recruited to classify the subjects between normal and deficient groups. In contrast to the parameter A4 mentioned before, the parameter A5 was demonstrated in [12] that it was also useful for the classification between qi-vacuity and ying-vacuity patients.

Table 9. The accuracy of discrimination between normal and qi-vacuity groups.

| Voice parameters | {A1, A2, A3} | {A1, A2} | A1 | A2 | A3 | A5 |
|---|---|---|---|---|---|---|
| Accuracy | 95.2% | 90.4% | 81.0% | 81.0% | 71.4% | 66.7% |

Table 10. The accuracy of discrimination between ying-vacuity and qi-vacuity groups.

| Voice parameters | A5 | {A4, A5} | {A5, A1} | {A5, A3} | {A3, A4, A5} |
|---|---|---|---|---|---|
| Accuracy | 80.0% | 80.0% | 76.0% | 76.0% | 64.3% |

Table 11. The accuracy of discrimination between normal and deficient groups.

| Voice parameters | A1 | {A1,A4} | {A1, A2} | {A1, A2,A3} | {A1, A3} |
|---|---|---|---|---|---|
| Accuracy | 81.1% | 78.4% | 75.7% | 75.7% | 75.7% |

## Discussion

In this research, we found the following results revealed in all subjects: (1) Although the same sentence was read by each subject three times, the distributions of FD values each time were not exactly the same; (2) No significant correlation was found between speech variations and syllables, it was various case by case; (3) The utterances of test sentences of both qi-vacuity and ying-vacuity patients have much more variations than that of normal subjects. Any malfunction of an organ would change the pattern of voice, which is detectable through diagnosis with the fractal dimension analysis method. It may be due to the physiological or mental problems. There are indeed some speech problems in stability and the ability for some test subjects to repeat the same sentence; (4) Good classification results were obtained using both sets of test sentences for the classification between deficient and normal subjects; (5) The classification accuracy using FD values of continuous speech was better than the results using five acoustic parameters of vowel /a/ in previous study. The results show that the objective auscultation in TCM by the variations of continuous speech are feasible and much more reliable than that by the variations of vowels.

## Conclusions

Although it had great achievement to modernize the auscultation in traditional Chinese medical diagnosis, most of researches emphasized to evaluate vowels in their experiments and they were not the only judgments of auscultative diagnosis. Speech and sounds uttered by humans can reflect important physiological information. For this reason, this study used the variations of continuous speech to distinguish between deficient patients and normal subjects. The utterances of test sentences of both qi-vacuity and ying-vacuity patients had much more variations than that of normal subjects. Deficient patients had malfunction of an organ would change the pattern of voice. As result of physiological or mental problems, there are certain speech problems in stability and the ability for some test subjects to repeat the same sentence. By these results of variations of continuous speech, we might conclude that the health condition of human body could be represented in the voice, and different situation (qi-vacuity and ying-vacuity) would cause different variations in voice. Also, the fractal dimension analysis of continuous speech provides an alternative and outstanding approach to the auscultation of Chinese medical diagnosis. However, this study is still lack of subjects, it requires additional investigation in the future work.

## Acknowledgments

The authors would like to thank the National Science Council, R.O.C., for partially supporting this work under grant number NSC-91-2320-B-035-001. Additionally, the authors also wish to thank Dr. Chung-Hsien Yang at the China Medical College Hospital, Taichung, Taiwan, R.O.C., for his counsel in regard to the work presented herein.

## References

[1] C.C. Chiu, H.H. Chang, and C.H. Yang, "Objective auscultation for traditional Chinese medical diagnosis using novel acoustic parameters", *Computer Methods and Programs in Biomedicine*, 62(2): 99-107, 2000.

[2] E. Rontal, M. Rontal, and M.I. Rolnick, "Objective evaluation of vocal pathology using voice spectrography", *Readings in Clinical Spectrography of Speech, Singular Publishing Group, San Diego, California*, 554-563, 1991.

[3] A.P. Accardo and E. Mumolo, "An algorithm for automatic differentiation between the speech of normals and patients with Fridreich's ataxia based on the short-time fractal dimension", *Computers in Biology and Medicine*, 28: 75-89, 1998.

[4] N. Wiseman and Y. Feng, A Practical Dictionary of Chinese Medicine, Paradigm Publications, Brookline, Massachusetts, 486, 1998.

[5] R.J.B. Hemler, G.H. Wieneke, and P.H. Dejonckere, "The effect of relative humidity of inhaled air on acoustic parameters of voice in normal subjects", *Journal of Voice*, 11: 295-300, 1997.

[6] C.H. Horng, The principles and methods of diagnostics, in The Illustrations of Chinese Medicine, Ch.4, Lead Press, Taipei, Taiwan, 1993.

[7] S.M. Mo and K.S. Tsai, "Clinical research of objective Chinese medical voice diagnosis", *Journal of China Fundamental Medicine of Traditional Chinese Medicine*, 4(5): 37-43, 1998.

[8] S.L. Wang, "Voice analysis of cough sound from Stage III Pulmonary TB patients", *Journal of Hu-Nan Medical College of Chinese Medicine*, 17(4): 33-36, 1997.

[9] C.S. Jwo, The Investigation of Novel Acoustic Parameters for Listening Examination of Chinese Medicine, Master Thesis, Institute of Automatic Control Engineering, Feng Chia University, Taiwan, 1998.

[10] C.S. Weng, L.Y. Shyu, Y.H. Chang, and L.N. Chiang, "A study of acoustical method in traditional Chinese medicine for the upper GI disorder patients", *Chinese Journal of Medical and Biological Engineering*, 18(4): 245-252, 1998.

[11] N.F. Shao and Q.L. Wang, Diagnostics of Traditional Chinese Medicine, Shandong Science and Technology Press, 1990.

[12] M.T. Yang, Using Fractal Dimension Analysis on Auscultation of Continuous Speech, Master Thesis, Institute of Automatic Control Engineering, Feng Chia University, Taiwan, 2001.

[13] B.B. Mandelbort, The Fractal Geometry of Nature, Freeman, New York, 1982.

[14] P. Maragos, "Measuring the fractal dimension of signals: morphological covers and iterative optimization", *IEEE Transactions On Signal Processing*, 41: 108-121, 1993.

[15] M.F. Barnsley, Fractals Everywhere, Academic Press, San Diego, 1988.

[16] A. Penn, "Estimating fractal dimension with fractal interpolation models", *IEEE Transactions on Medical Imaging*, 16(6): 930-937, 1997.

[17] H.F.V. Boshoff, "A fast box counting algorithm for determining the fractal dimension of sampled continuous functions", *Proce. of the 1992 South America Symposium of Communications and Signal Processing*, 43-48, 1992.

[18] L. Rabiner, B.H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[19] H. Sakoe, S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26: 43-49, 1978.

[20] A.K. Jain, R.C. Dubes, Algorithms for Clustering Data, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

# 利用碎形維度分析方法於中醫聞診現代化研究

邱創乾* 楊明達 林宸生

逢甲大學自動控制工程研究所

收件日期 2002 年 10 月 17 日；接受日期 2002 年 11 月 8 日

## 摘 要

中醫聞診以患者聲音的變化爲診斷的依據，是臨床辨證重要的一環。本研究採用連續發音的方式探討十六名氣虛患者，九名陰虛患者及十二名正常人在連續語音表現上的差異性。我們選用碎形維度演算法來計算這類非規律性語音信號所表現出的特性，所採用的測試語句共有兩組，一爲特定語音設計句，另一爲隨意選取句。研究方法是讓測試者連續發同一句語音三次，觀察其碎形維度在這三筆之間複雜程度的差異性，以動態時間扭曲路徑演算法，來量化同一語句中三筆語音之間的型態差異。結果顯示：(1)雖然每一位測試者唸出相同的句子，但是彼此間的碎形維度值分佈變化並不一定相似；(2)由兩組不同的測試句子可觀察得到若比較同一人前後三筆碎形維度值，語音變化與音節種類並無太大關係，而是隨著病人不同而改變；(3)氣虛與陰虛患者在重複三次測試句的差異變化程度比正常人要大許多；(4)兩組測試句同時用於分類正常人與虛症患者上，可得到很好的分類結果(85%以上)；(5)利用碎形維度參數於連續語音的分類結果其準確率比起過去研究所使用的五個語音參數於元音/a/之分類結果要來得好。因此連續音之碎形維度分析可提供另一種優異的工具於中醫聞診現代化之研究上。

**關鍵詞**：聞診、中醫、碎形維度、動態時間扭曲

* 通訊作者：邱創乾
電話：+886-4-24517250 轉 3930；傳真：+886-4-24515701
電子郵件信箱：chiuc@auto.fcu.edu.tw